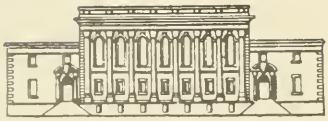
SWEET BRIAR COLLEGE
3 2449 0312734 %

Presented by
Rapti Manohara De Silva '88

SWEET
BRIAR
COLLEGE
LIBRARY



Digitized by the Internet Archive
in 2010 with funding from
Lyrasis Members and Sloan Foundation

http://www.archive.org/details/twospectraltheor00desi

Rapti de Silva

April 6th, 1988

Senior Honours

Professor Kirkwood

TWO SPECTRAL THEOREMS

THE JORDAN CANONICAL FORM
FOR LINEAR OPERATORS
AND
THE SPECTRAL THEOREM
FOR NORMAL OPERATORS

A Thesis
Submitted in Partial Fulfillment of the Requirements
For the Honours degree in Mathematics.
Sweet Briar, April 1988.

Archives
QA
184
.D47
1988

To Mr. Kirkwood

With Many Thanks

CONTENTS

Preface

The Jordan Normal Form for Linear Operators

Introduction 1
Some Basic Definitions 2
The Matrix Representation of a Linear Operator 3
Bases for the Matrix of a Linear Operator 6
The Diagonal Form 8
The Direct Sum Decomposition of a Vector Space:
Invariant Subspaces and the
Reduction of a Linear Operator 10
Projections 14
The Minimum Polynomial of a Linear Operator 25
Eigenvalues and Eigenvectors 27
The Characteristic Polynomial of a Linear Operator 33
The Decomposition of an Operator 36
The Special Basis for the Jordan Form of
a Nilpotent Operator 41
The Jordan Form 53

The Spectral Theorem for Normal Operators

Introduction 62
The Basic Definitions and Theorems 63
The Spectral Theorem for Normal Operators 65

Conclusion 72

Notes 73

Bibliography 74

Preface

The term "spectral theorem" comes from the word spectrum which means any range of characteristics, values, activities, etc. Thus a spectral theorem for an operator describes how the operator is determined by its spectrum. A spectral theorem is the single most powerful technique for describing the structure of an operator on a vector space.

The Jordan Normal Form

for Linear Operators

Introduction

The object of this chapter will be to develop a simple representation of a linear operator (or linear transformation) on a finite-dimensional vector space over an algebraically closed field. For computational purposes, this representation is a matrix.

The Jordan Normal (Canonical) Form of a matrix gives an extremely simple representation of a linear operator on a complex finite-dimensional vector space. (The complex vector space ensures that the field is algebraically closed). We emphasize that the discussion is limited to finite-dimensional spaces since there is no analogue for the infinite-dimensional case. By studying certain invariants of the linear operator such as invariant subspaces, projections, eigenvalues, and eigenvectors, we shall examine how a specific operator is determined. Based on this determination, a special basis will be developed for the given vector space, with respect to which the matrix of the operator on that vector space is quite simple. This matrix will be the Jordan Normal Form.

SOME BASIC DEFINITIONS

We will first define some of the concepts necessary for pursuing our objective.

A **linear operator** $T : U \dashrightarrow V$, where U and V are vector spaces over the same scalar field, is a function such that

i) $T(x + y) = T(x) + T(y)$ for all $x, y \in U$

ii) $T(\alpha x) = \alpha T(x)$ for all $x \in U$ and scalars α

We say that T is a **linear operator on U** if $T : U \dashrightarrow U$.

A **basis** for V is a collection of vectors $\{x_1, x_2, \ldots, x_n\}$ in V such that

i) $\{x_1, x_2, \ldots, x_n\}$ is linearly independent.

i.e; if $\alpha_1 x_1 + \alpha_2 x_2 + \ldots + \alpha_n x_n = 0$, then $\alpha_1 = \alpha_2 = \ldots = \alpha_n = 0$

ii) $\{x_1, x_2, \ldots, x_n\}$ spans V.

The **range space** of $T : U \dashrightarrow V$, $\mathbf{R_T}$, is defined by

$$R_T = \{ y \mid y \in V \text{ and } y = T(x) \text{ for some } x \in U \}$$

The **null space** of $T : U \dashrightarrow V$, $\mathbf{N_T}$, is defined by

$$N_T = \{ x \mid x \in U \text{ and } T(x) = 0 \}$$

THE MATRIX REPRESENTATION OF A LINEAR OPERATOR

A linear operator is most conveniently represented as a matrix. It is important to realize that while a linear operator is a basis-free object, a particular matrix representation of $\mathcal{T} : U \dashrightarrow V$, depends on the choice of bases for U and V. Thus if we know the effect of $\mathcal{T}$ on the basis vectors, we can determine its effect on any vector.

It is known that any vector space over the field $\mathbb{R}$ with basis $\mathcal{X} = \{x_1, x_2, \ldots, x_n\}$ is **isomorphic** to $\mathbb{R}^n$ with the isomorphism being a mapping into a column vector as follows:

$$\Phi: \quad \beta_1 x_1 + \beta_2 x_2 + \ldots + \beta_n x_n \quad \dashrightarrow \quad \begin{bmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \vdots \\ \beta_n \end{bmatrix}$$

In order to simplify notation, we will henceforth identify the vector $\beta_1 x_1 + \beta_2 x_2 + \ldots + \beta_n x_n$ with the column vector given above.

Theorem 1

A linear operator $\mathcal{T} : U \dashrightarrow V$ is uniquely represented by a matrix $\mathcal{A}_{\mathcal{T}}$, where $\mathcal{A}_{\mathcal{T}}$ is determined by $\mathcal{T}$ and the choice of bases for both U and V.

Proof

Assume $\mathcal{X} = \{x_1, x_2, \ldots, x_n\}$ and $\mathcal{Y} = \{y_1, y_2, \ldots, y_m\}$ are bases for U and V respectively.

Since $\mathcal{Y}$ is a basis for V and given $T(x_j) \in V$, there exists scalars α_{ij} such that

$$T(x_j) = \sum_{i=1}^{m} \alpha_{ij} y_i$$

For $x \in U$, since $\mathcal{X}$ is a basis for U, there exist scalars β_j such that

$$x = \sum_{j=1}^{n} \beta_j x_j$$

Then since T is linear,

$$T(x) = T\left(\sum_{j=1}^{n} \beta_j x_j\right) = \sum_{j=1}^{n} \beta_j \left(T(x_j)\right) = \sum_{j=1}^{n} \beta_j \left(\sum_{i=1}^{m} \alpha_{ij} y_i\right)$$

$$= \sum_{j=1}^{n} \beta_j \left(\alpha_{1j} y_1 + \alpha_{2j} y_2 + \ldots + \alpha_{mj} y_m\right)$$

$$= \left(\sum_{j=1}^{n} \beta_j \alpha_{1j}\right) y_1 + \left(\sum_{j=1}^{n} \beta_j \alpha_{2j}\right) y_2 + \ldots + \left(\sum_{j=1}^{n} \beta_j \alpha_{mj}\right) y_m$$

$$T(x) = \begin{bmatrix} \sum_{j=1}^{n} \beta_j \alpha_{1j} \\ \vdots \\ \vdots \\ \sum_{j=1}^{n} \beta_j \alpha_{mj} \end{bmatrix}$$

We emphasize that when summed over $j = 1, 2, \ldots, n$, the resulting vector $T(x)$ is obtained with respect to both the bases $\mathcal{X}$ and $\mathcal{Y}$.

We now expand the above column vector representation of $\mathcal{T}(x)$ as follows:

$$\mathcal{T}(x) = \begin{bmatrix} \beta_1\alpha_{11} + \beta_2\alpha_{12} + \dots + \beta_n\alpha_{1n} \\ \beta_1\alpha_{21} + \dots : \dots + \beta_n\alpha_{2n} \\ \vdots \qquad \vdots \qquad \vdots \\ \vdots \qquad \vdots \qquad \vdots \\ \beta_1\alpha_{m1} + \beta_2\alpha_{m2} + \dots + \beta_n\alpha_{mn} \end{bmatrix}$$

For computational purposes the vector $\mathcal{T}(x)$ may be expressed as the following matrix product:

$$\mathcal{T}(x) = \begin{bmatrix} \alpha_{11} & \alpha_{12} \dots\dots & \alpha_{1n} \\ \alpha_{21} & \vdots \dots\dots & \alpha_{2n} \\ \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots \\ \alpha_{m1} & \alpha_{m2} \dots\dots & \alpha_{mn} \end{bmatrix} \begin{bmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \vdots \\ \vdots \\ \beta_n \end{bmatrix}$$

But $\begin{bmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_n \end{bmatrix}$ is the vector x with respect to the basis $\mathcal{X}$.

$$\text{Therefore} \quad \mathcal{T}(x) = \mathcal{A}_{\mathcal{T}}\, x$$

where $\mathcal{A}_{\mathcal{T}} = (\alpha_{ij})$ is the m x n matrix with α_{ij} as the entry in the i-th row and j-th column of the matrix. We again emphasize that α_{ij} is determined by both the operator and the choice of bases ■

Bases for the matrix of a linear operator

As stated at the outset, in developing the Jordan form of a matrix, our object is to identify the "preferred basis" of a matrix. From now on, we will restrict out attention to the case where $\mathcal{T}$ is a linear operator on V. Thus the matrix $\mathcal{A}_{\mathcal{T}} = (\alpha_{ij})$ is a square array. Identifying the preferred basis is equivalent to identifying the preferred form of the matrix $\mathcal{A}_{\mathcal{T}}$.

We will first consider how the matrix of an operator changes with a change of basis.

Proposition

Let $\mathcal{T}$ be a linear operator on the n-dimensional vector space V and let $\mathcal{X} = \{x_i\}_{i=1}^{n}$ and $\mathcal{Y} = \{y_i\}_{i=1}^{n}$ be two different bases for V. If

$\mathcal{A}_{\mathcal{T},\mathcal{X}} = (\alpha_{ij})$ is the matrix of $\mathcal{T}$ with respect to $\mathcal{X}$,

and $\mathcal{B}_{\mathcal{T},\mathcal{Y}} = (\beta_{ij})$ is the matrix of $\mathcal{T}$ with respect to $\mathcal{Y}$,

then there is a matrix $\mathcal{C}$ such that $\mathcal{C}\mathcal{A}_{\mathcal{T},\mathcal{X}} = \mathcal{B}_{\mathcal{T},\mathcal{Y}}\mathcal{C}$.

Proof

For each $x_j \in \mathcal{X}$, there are scalars α_{ij}, $i = 1, 2, \ldots, n$, such that

$$\mathcal{T}(x_j) = \sum_{i=1}^{n} \alpha_{ij} x_i$$

Let $y_j = \sum_{i=1}^{n} c_{ij} x_i$ give the change of coordinates from $\mathcal{Y}$ to $\mathcal{X}$. Then

$$T(y_j) = T \sum_{i=1}^{n} c_{ij} x_i = \sum_{i=1}^{n} c_{ij} T(x_i) = \sum_{i=1}^{n} c_{ij} \sum_{k=1}^{n} \alpha_{ki} x_k$$

Likewise for each $y_i \in \mathcal{Y}$, there are scalars β_{ij}, $i = 1, 2, \ldots, n$, such that

$$T(y_j) = \sum_{i=1}^{n} \beta_{ij} y_i = \sum_{i=1}^{n} \beta_{ij} \sum_{k=1}^{n} c_{ki} x_k$$

Therefore by equating the expressions for $T(y_j)$ we get,

$$\sum_{i=1}^{n} \sum_{k=1}^{n} c_{ij} \alpha_{ki} x_k = \sum_{i=1}^{n} \sum_{k=1}^{n} \beta_{ij} c_{ki} x_k$$

Let $\mathcal{C}$ be the matrix whose ij-th entry is c_{ij}. Then we obtain the equation

$$\mathcal{C} \mathcal{A}_{T,\mathcal{X}} = \mathcal{B}_{T,\mathcal{Y}} \mathcal{C} \quad \square$$

If $\mathcal{C}$ is invertible, then $\qquad \mathcal{A}_{T,\mathcal{X}} = \mathcal{C}^{-1} \mathcal{B}_{T,\mathcal{Y}} \mathcal{C}$

N. B. Although we shall not prove it here, it is true that $\mathcal{C}$ is invertible in the case where $\mathcal{A}_{T,\mathcal{X}}$ and $\mathcal{B}_{T,\mathcal{Y}}$ are both matrix representations of T. In fact, matrices $\mathcal{A}$ and $\mathcal{B}$ both represent the same linear operator with respect to different bases, if and only if there is an invertible matrix $\mathcal{C}$ with $\mathcal{A} = \mathcal{C}^{-1} \mathcal{B} \mathcal{C}$.

A matrix $\mathcal{A}$ is said to be **similar** to a matrix $\mathcal{B}$ if and only if there is an invertible matrix $\mathcal{C}$ such that $\mathcal{A} = \mathcal{C}^{-1} \mathcal{B} \mathcal{C}$. Thus in our discussion above,

$\mathcal{A}_{T,\mathcal{X}}$ is similar to $\mathcal{B}_{T,\mathcal{Y}}$. If $\mathcal{A}_{T,\mathcal{X}}$ is similar to yet another matrix $\mathcal{D}$, then $\mathcal{D}$ is simply the matrix of T with respect to another basis.

THE DIAGONAL FORM

The most convenient form of a matrix representation of a linear operator would be if the matrix were in diagonal form. We will now give an example where the matrix can be put into diagonal form and an example where it is impossible to do so.

Example 1

Let $T : \mathbb{R}^2 \dashrightarrow \mathbb{R}^2$ be defined by its operation on the basis $\mathcal{X} = \{(1,0),(0,1)\}$ as follows : $T(1,0) \dashrightarrow (1,3)$ and $T(0,1) \dashrightarrow (2,2)$. Then

$$\mathcal{A}_{T,\mathcal{X}} = \begin{vmatrix} 1 & 2 \\ 3 & 2 \end{vmatrix}$$

In order to show that T has a diagonal matrix representation, we must find a diagonal matrix $\mathcal{D}$ such that $\mathcal{A}_{T,\mathcal{X}}\, \mathcal{C} = \mathcal{C}\, \mathcal{D}$. Consider

$$\begin{bmatrix} 1 & 2 \\ 3 & 2 \end{bmatrix} \begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix}$$

Then by taking the matrix product and equating the corresponding matrix

entries we get $a + 2c = \lambda_1 a$ and $3a + 2c = \lambda_1 c$

Then $2a = \lambda_1 (c - a)$. Thus $\lambda_1 = 2a / (c - a)$.

Similarly, we find that $\lambda_2 = 2b / (d - b)$.

If we let $a = 2$, $b = 1$, $c = 1$, and $d = 3$, then $\det \mathbb{C} \neq 0$ and $\mathbb{C}$ is an invertible

matrix. Also $$\mathbb{D} = \begin{bmatrix} -4 & 0 \\ 0 & 1 \end{bmatrix}$$

Thus $\mathbb{T}$ has a diagonal matrix representation.

Example 2

Let $\mathbb{T} : \mathbb{R}^2 \dashrightarrow \mathbb{R}^2$ be the rotation of the basis vectors $\mathbb{X} = \{(1,0),(0,1)\}$ through an angle of $\pi/4$ in the counter-clockwise direction. Then

$$\mathbb{T}(1,0) \dashrightarrow (\operatorname{Cos} \pi/4, \operatorname{Sin} \pi/4) = (1/\sqrt{2}, 1/\sqrt{2})$$

$$\mathbb{T}(0,1) \dashrightarrow (-\operatorname{Sin} \pi/4, \operatorname{Cos} \pi/4) = (-1/\sqrt{2}, 1/\sqrt{2})$$

Thus

$$\mathcal{A}_{\mathbb{T},\mathbb{X}} = \begin{bmatrix} 1/\sqrt{2} & -1/\sqrt{2} \\ 1/\sqrt{2} & 1/\sqrt{2} \end{bmatrix}$$

As in Example 1, we will try to find matrices $\mathbb{C}$ and $\mathbb{D}$ such that $\mathcal{A}_{\mathbb{T},\mathbb{X}}\, \mathbb{C} = \mathbb{C}\, \mathbb{D}$ where $\mathbb{C}$ is an invertible matrix and $\mathbb{D}$ is diagonal.

$$\begin{bmatrix} 1/\sqrt{2} & -1/\sqrt{2} \\ 1/\sqrt{2} & 1/\sqrt{2} \end{bmatrix} \begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix}$$

By equating corresponding matrix entries as before, we get

$$a - c = \sqrt{2}\, a\, \lambda_1 \quad \text{and} \quad a + c = \sqrt{2}\, c\, \lambda_1$$

Then $\quad (a - c)/(a + c) = a/c \quad$ and $\quad a^2 + c^2 = 0$

Thus $\quad a = 0 \quad$ and $\quad c = 0$

But then det $\mathbf{C} = 0$. Hence $\mathbf{C}$ is not invertible. Thus there is no diagonal matrix representation of T.

The Direct Sum Decomposition of a Vector Space : Invariant Subspaces and the Reduction of a Linear Operator

Since the diagonal matrix fails to represent all linear operators, we continue our analysis of the invariant properties of linear operators. We shall, however, see the concept of the "diagonal" return in our examination of the effect a linear operator has on certain subspaces of a vector space.

Definition

A subspace U of V is **invariant** under $T : V \dashrightarrow V$ if $x \in U$ implies that $T(x) \in U$.

The importance of U being an invariant subspace of T is that T

(restricted to U) is in fact a linear operator on U.

We shall now continue to simplify the matrix representation of $\mathcal{T}$ by distinguishing smaller subspaces of V as invariant subspaces under restrictions of the linear operator $\mathcal{T}$, and by identifying preferred bases for each invariant subspace.

Let $\mathcal{T}$ be a linear operator on the n-dimensional vector space V, and let U be a k-dimensional subspace of V that is invariant under $\mathcal{T}$.

Suppose $\mathcal{X}_1 = \{x_1, x_2, \ldots, x_k\}$ is a basis for U. We can extend this basis to a basis $\mathcal{X} = \{x_1, x_2, \ldots, x_k, x_{k+1}, \ldots, x_n\}$ of V. Notice that since for any vector $u \in U$, $\mathcal{T}(u) \in U$, then

$$
\begin{aligned}
\mathcal{T}(x_1) &= \alpha_{11}x_1 + \alpha_{12}x_2 + \ldots + \alpha_{1k}x_k + 0x_{k+1} + \ldots + 0x_n \\
\mathcal{T}(x_2) &= \alpha_{21}x_1 + \alpha_{22}x_2 + \ldots + \alpha_{2k}x_k + 0x_{k+1} + \ldots + 0x_n \\
&\vdots \\
\mathcal{T}(x_k) &= \alpha_{k1}x_1 + \alpha_{k2}x_2 + \ldots + \alpha_{kk}x_k + 0x_{k+1} + \ldots + 0x_n \\
\mathcal{T}(x_{k+1}) &= \alpha_{k+11}x_1 + \alpha_{k+12}x_2 + \ldots\ldots\ldots\ldots + \alpha_{k+1n}x_n \\
\mathcal{T}(x_{k+2}) &= \alpha_{k+21}x_1 + \alpha_{k+22}x_2 + \ldots\ldots\ldots\ldots + \alpha_{k+2n}x_n \\
&\vdots \\
\mathcal{T}(x_n) &= \alpha_{n1}x_1 + \alpha_{n2}x_2 + \ldots\ldots\ldots\ldots + \alpha_{nn}x_n
\end{aligned}
$$

Thus with respect to the basis $\mathcal{X}$, the matrix representation of $\mathcal{T}$ will have

the form

$$\mathcal{A}_{\mathcal{T}} = \begin{bmatrix} T_1 & T_3 \\ 0 & T_2 \end{bmatrix}$$

where $[T_1]$ is the $k \times k$ matrix representation of the $\mathcal{T}$ restricted to U.

$[T_2]$ is a $(n-k) \times (n-k)$ matrix.

and $[T_3]$ is a $k \times (n-k)$ matrix.

If $[T_3] = 0$, we would get the following "diagonal" form:

$$\mathcal{A}_{\mathcal{T}} = \begin{bmatrix} T_1 & 0 \\ 0 & T_2 \end{bmatrix}$$

Notice that $[T_3] = 0$, if and only if $U^c \supset \mathcal{T}(U^c)$. In order to obtain the above form, we now identify a special set of subspaces.

Two subspaces U and W of a vector space V are said to be **complementary** (or the **complement** of each other) if $U \cap W = \{0\}$ and $U + W = V$.

V is the **direct sum** of U and W if and only if U and W are complementary subspaces of V. In this case we use the notation $V = U \oplus W$.

We shall use the following result in our discussion:

Lemma

$V = U \oplus W$ if and only if each $z \in V$ can be written uniquely as

$z = x + y$ where $x \in U$ and $y \in V$ □

Definition

Let $T : V \dashrightarrow V$ and let $V = U \oplus W$. If U and W are both invariant under T, then T is said to be **reduced** by the pair (U,W).

The importance of the reduction of a linear operator T by (U,W) is that T can be regarded as the direct sum of two operators: T_1, the restriction of T to U; and T_2, the restriction of T to W. This follows from the fact that for $z \in V$ and $z = x + y$,

$$T(z) = T(x + y) = T(x) + T(y) = T_1(x) + T_2(y) \quad \text{(since } x \in U \text{ and } y \in V\text{)}$$

Thus we shall write $$T = T_1 \oplus T_2$$

We extend the idea of the reduction of a linear operator as follows: Let $T : V \dashrightarrow V$ and $V = U_1 \oplus U_2 \oplus \ldots \oplus U_k$. Let T be reduced by $(U_1, U_2, \ldots, U_k)$, and let T_i be the restriction of T to U_i. Then $T = T_1 \oplus T_2 \oplus \ldots \oplus T_k$.

We now modify our original bases $\mathcal{X}$ and $\mathcal{X}_1$ to accomodate the reduction of the linear operator T on V by (U,W), the complementary subspaces of V which are both invariant under T.

Given $\mathcal{X}_1 = \{x_1, x_2, \ldots, x_k\}$ as a basis for U and $\mathcal{X}_2 = \{x_{k+1}, x_{k+2}, \ldots, x_n\}$ as a basis for W, if $V = U \oplus W$, then $\mathcal{X} = \{x_1, x_2, \ldots, x_n\}$ is a basis for V.

Then since U is invariant under $\mathcal{T}$,

$$\mathcal{T}(x_j) = \sum_{i=1}^{k} \alpha_{ij} x_i \qquad j = 1, 2, \ldots, k$$

and since W is invariant under $\mathcal{T}$,

$$\mathcal{T}(x_j) = \sum_{i=k+1}^{n} \alpha_{ij} x_i \qquad j = k+1, k+2, \ldots, n$$

Thus the matrix of $\mathcal{T}$ with respect to the basis $\mathcal{X}$ will be of the "diagonal" form

$$\mathcal{A}_{\mathcal{T}} = \begin{bmatrix} T_1 & 0 \\ 0 & T_2 \end{bmatrix}$$

where $[T_1]$ is the matrix of $\mathcal{T}_1$ (the restriction of $\mathcal{T}$ to U) with respect to the basis $\mathcal{X}_1$, and $[T_2]$ is the matrix of $\mathcal{T}_2$ (the restriction of $\mathcal{T}$ to W) with respect to the basis $\mathcal{X}_2$.

Projections

In discussing the **direct sum decomposition** of the vector space V into U and W, each an invariant subspace under $\mathcal{T}$, we examined the special case where the k x (n - k) matrix $[T_3] = 0$.

We now describe a special linear operator $\mathcal{E}$ that helps distinguish the invariant subspaces under the linear operator $\mathcal{T}$, and in doing so reduces $\mathcal{T}$.

Definition

The operator $\mathcal{E}$ on V is a **projection** on V if

i) $V = U \oplus W$ where $U = R_{\mathcal{E}}$ and $W = N_{\mathcal{E}}$

ii) If for any $z \in V$, $\mathcal{E}(z) = x$ where $z = x + y$ with $x \in U$, $y \in W$

We say that $\mathcal{E}$ is a projection on U ($R_{\mathcal{E}}$) along W ($N_{\mathcal{E}}$).

We first note that a projection is indeed a linear operator. Given conditions i) and ii), let $z_1, z_2 \in V$ where $z_1 = x_1 + y_1$ and $z_2 = x_2 + y_2$, and let α be any scalar.

Then $$(z_1 + z_2) = (x_1 + x_2) + (y_1 + y_2)$$

Since $(x_1 + x_2) \in U$ and $(y_1 + y_2) \in W$,

from the definition $$\mathcal{E}(z_1 + z_2) = (x_1 + x_2) = \mathcal{E}(z_1) + \mathcal{E}(z_2)$$

On the other hand, $$\alpha z_1 = \alpha x_1 + \alpha y_1$$

Now since $\alpha x_1 \in U$ and $\alpha y_1 \in W$, $\mathcal{E}(\alpha z_1) = \alpha x_1 = \alpha \mathcal{E}(z_1)$

Thus $\mathcal{E}$ is linear. Since $\mathcal{E} : V \dashrightarrow U$ and $V \supset U$, $\mathcal{E}$ is a linear operator on V $\square$

We emphasize an important connection between projections and the direct sum decomposition of a vector space. Given a projection on V, we automatically create a direct sum decomposition of V; i.e. $V = R_{\mathcal{E}} \oplus N_{\mathcal{E}}$.

Conversely, any direct sum decomposition of V induces a projection operator.

The following theorem gives some properties of projections that we shall use later.

Theorem 2

Let $\mathcal{E}$ be a linear operator on V. Then

i) $\mathcal{E}$ is a projection on V if and only if $\mathcal{E}^2 = \mathcal{E}$

ii) $\mathcal{E}$ is a projection on a subspace U along a subspace W if and only if $I - \mathcal{E}$ is a projection on W along U.

Proof

(i) Let us assume that $\mathcal{E}$ is a projection on V. We must show that $\mathcal{E}^2 = \mathcal{E}$, that is, $\mathcal{E}^2(z) = \mathcal{E}(z)$ for any $z \in V$. Since $\mathcal{E}$ is a projection on V, we know that $V = R_{\mathcal{E}} \oplus N_{\mathcal{E}}$ and given any $z \in V$,

we can write $\quad z = x + y \quad$ where $x \in R_{\mathcal{E}}$, $y \in N_{\mathcal{E}}$

Then

$$\begin{aligned}\mathcal{E}^2(z) &= \mathcal{E}(\mathcal{E}(z)) \\ &= \mathcal{E}(\mathcal{E}(x+y)) \\ &= \mathcal{E}(\mathcal{E}(x) + \mathcal{E}(y)) \\ &= \mathcal{E}(\mathcal{E}(x)) \\ &= \mathcal{E}(x)\end{aligned}$$

But

$$\begin{aligned}\mathcal{E}(z) &= \mathcal{E}(x+y) \\ &= \mathcal{E}(x)\end{aligned}$$

Therefore $\quad \mathcal{E}^2 = \mathcal{E}$.

Conversely, suppose $\mathcal{E}^2 = \mathcal{E}$. We must show that $\mathcal{E}$ is a projection on V.

Let $\quad W = \{ y \in V \mid \mathcal{E}(y) = 0 \}$ and $U = \{ x \in V \mid \mathcal{E}(x) = x \}$

Then $\quad W = N_{\mathcal{E}} \quad$ and $\quad R_{\mathcal{E}} \supset U$

Now let $x \in R_{\mathcal{E}}$. Then there exists a $z \in V$ such that

$$\mathcal{E}(z) = x$$

Then $\quad \mathcal{E}^2(z) = \mathcal{E}(x)$

But since $\mathcal{E}^2 = \mathcal{E}$, $\quad \mathcal{E}^2(z) = \mathcal{E}(z) = x$

Therefore $\quad \mathcal{E}(x) = x$

Thus $x \in U$ and $U \supset R_{\mathcal{E}}$. $\quad \therefore \quad U = R_{\mathcal{E}}$

To show $\mathcal{E}$ is a projection on V it now suffices to show that $R_{\mathcal{E}} \cap N_{\mathcal{E}} = \{0\}$.

Let $x \in R_{\mathcal{E}} \cap N_{\mathcal{E}}$. Then $\quad x \in R_{\mathcal{E}} \quad$ and $\quad x \in N_{\mathcal{E}}$

$$\mathcal{E}(x) = x \quad \text{and} \quad \mathcal{E}(x) = 0$$

$$\therefore \quad x = 0$$

Therefore $\quad \mathcal{E}$ is a projection on V.

ii) Let us now assume that $\mathcal{E}$ is a projection on U along W and show that $I - \mathcal{E}$ is a projection on W along U. First consider $I - \mathcal{E}$:

$$(I - \mathcal{E})^2 = (I - \mathcal{E})(I - \mathcal{E}) = I\,I - I\,\mathcal{E} - \mathcal{E}I + \mathcal{E}\,\mathcal{E} = I^2 - \mathcal{E} - \mathcal{E} + \mathcal{E}^2 = I - 2\,\mathcal{E} + \mathcal{E} = I - \mathcal{E}$$

Thus $I - \mathcal{E}$ is a projection. We show that $I - \mathcal{E}$ is a projection on W along U.

Now given any $z \in V$, since $E^2 = E$,

$$(E - E^2)\, z = 0$$

Then $$E\,[\,(I - E)\, z\,] = 0$$

But if $E(x) = 0$, $x = 0$ or $x \in N_E = W$. Thus $(I - E)\, z = 0$ or $(I - E)\, z \in W$ for all $z \in V$. Therefore $(I - E)$ is a projection on W along U.

Conversely, let us suppose that $(I - E)$ is a projection on W along U. We must show that E is a projection on U along W.

Since $(I - E)$ is a projection,

$$(I - E)^2 = I - E$$

Then $$(I - E)^2 - (I - E) = 0$$

$$I^2 - 2E + E^2 - I + E = 0$$

$$I - E + E^2 - I = 0$$

$$\therefore \quad E^2 = E\,.$$

Therefore E is a projection.

Now since $(I - E)$ is a projection on $W = R_{I-E}$ along $U = N_{I-E}$, for any $z \in V$,

we can write $$z = x + y \qquad x \in R_{I-E},\ y \in N_{I-E}$$

Then $$(I - E)\, z = y = z - x$$

or $$z - E(z) = z - x$$

Thus for all $z \in V$, $$E(z) = x \in U$$

But E is a projection. Therefore E is a projection on U along W ■

The next two theorems show how a projection E helps distinguish invariant subspaces under the linear operator T and how E reduces T.

Theorem 5

Let T be a linear operator on V and let U be a subspace of V.

i) If U is invariant under T, then for every projection E on U, $ETE = TE$.

ii) If $ETE = TE$ for some projection E on U, then U is invariant under T.

Proof

i) Let us suppose that U is invariant under T and let E be any projection on U. We must show that $ETE = TE$.

Since E is a projection on U,

$$V = U \oplus W \qquad\qquad U = R_E,\ W = N_E$$

and for any $z \in V$, we can find a unique $x \in U$ and $y \in W$ such that $z = x + y$

and $$E(z) = x$$

Then since $x \in U$ and U is invariant under T,

$$TE(z) = T(x) = x$$

Also since $x \in U$ and E is a projection on U,

$$ETE(z) = E(x) = x$$

By equating the two expressions for x,

$$ETE = TE \qquad \text{for every projection } E \text{ on U.}$$

ii) Now let us assume that $ETE = TE$ for some projection E on U. We must show that U is invariant under T.

Let $x \in U$. We must show that $T(x) \in U$.

Since $ETE = TE$ for E a projection on U,

$$ETE(x) = TE(x)$$

Since $x \in U$ and E is a projection on U, $E(x) = x$. Thus

$$ET(x) = T(x)$$

But because E is a projection on U, this means that

$$T(x) \in U$$

Therefore U is invariant under T ■

Theorem 4

Let T be a linear operator on V and let $V = U \oplus W$. T is reduced by (U,W) if and only if $ET = TE$ for any projection E on U along W.

Proof

Let us suppose that T is reduced by (U,W) and that E is any projection on U along W. We must show that $ET = TE$.

Let $x \in U$. Since T is reduced by U, U is invariant under T and

$$T(x) \in U$$

Thus for any projection E on U,

$$ET(x) = T(x)$$

Also since E is a projection on U, $x = E(x)$. Then by substitution we get

$$ET(x) = TE(x)$$

Therefore $\mathcal{E}\mathcal{T} = \mathcal{T}\mathcal{E}$ for any projection $\mathcal{E}$ on U along W.

Conversely, let us suppose that $V = U \oplus W$ and for any projection $\mathcal{E}$ on U along W, $\mathcal{E}\mathcal{T} = \mathcal{T}\mathcal{E}$. We must show that (U,W) is a reducing pair for $\mathcal{T}$.

Since $V = U \oplus W$, it suffices to show that U and W are each invariant under $\mathcal{T}$.

Let $x \in U$. Then since $\mathcal{E}\mathcal{T} = \mathcal{T}\mathcal{E}$,

$$\mathcal{E}\mathcal{T}(x) = \mathcal{T}\mathcal{E}(x)$$

Also for $\mathcal{E}$ a projection on U and $x \in U$, $\mathcal{E}(x) = x$.

Thus

$$\mathcal{E}\mathcal{T}(x) = \mathcal{T}(x)$$

But this is possible only if $\mathcal{T}(x) \in U$. Therefore U is invariant under $\mathcal{T}$.

Now let $y \in W$. Then since $\mathcal{E}\mathcal{T} = \mathcal{T}\mathcal{E}$,

$$\mathcal{E}\mathcal{T}(y) = \mathcal{T}\mathcal{E}(y)$$

But since $W = N_{\mathcal{E}}$, for $y \in W$, $\mathcal{E}(y) = 0$.

Thus

$$\mathcal{E}\mathcal{T}(y) = 0$$

But this is possible only if $\mathcal{T}(y) \in W = N_{\mathcal{E}}$. Therefore W is invariant under $\mathcal{T}$ ■

Before we begin the next stage of developing the Jordan form, we give a result which emphasizes a more complete connection between the direct sum decomposition of a vector space and projections.

Theorem 5

If $E_1, E_2, \ldots, E_k$ are linear operators on a vector space V such that

a) $I = E_1 + E_2 + \ldots + E_k$ and b) $E_i E_j = 0$ if $i \neq j$, then

i) each E_i, $i = 1, 2, \ldots, k$, is a projection

ii) $V = W_1 \oplus W_2 \oplus \ldots \oplus W_k$ where $W_i = R_{E_i}$, $i = 1, 2, \ldots, k$

Conversely, if $V = W_1 \oplus W_2 \oplus \ldots \oplus W_k$, then

i) there exist projections $E_1, E_2, \ldots, E_k$ which satisfy conditions a) and b)

ii) $W_i = R_{E_i}$, $i = 1, 2, \ldots, k$

Proof

Suppose $E_1, E_2, \ldots, E_k$ are linear operators on V that satisfy the conditions a) and b).

Let a) be multiplied by E_i to get

$$E_i = E_i E_1 + E_i E_2 + \ldots + E_i E_i + \ldots + E_i E_k$$

Using condition b) this simplifies to give

$$E_i = E_i^2 \qquad i = 1, 2, \ldots, k$$

Therefore i) E_i is a projection for $i = 1, 2, \ldots, k$.

Now let $W_i = R_{E_i}$, $i = 1, 2, \ldots, k$. Then given $x \in V$, $E_i(x) \in W_i$. Let a) be applied to x to give

$$x = (E_1 + E_2 + \ldots + E_k)\, x$$

Thus $$x = \sum_{i=1}^{k} E_i(x) = \sum_{i=1}^{k} x_i \qquad \text{where } x_i \in W_i$$

This means that

$$V = W_1 + W_2 + \ldots + W_k$$

To show that V is the direct sum of the W_i s, it is sufficient to show that

for any $x \in V$, the above representation of x is unique.

Let $$x = \sum_{i=1}^{k} x_i = \sum_{i=1}^{k} y_i \qquad \text{where } x_i, y_i \in W_i$$

Then since E_i is a projection on W_i, $E_i(x_i) = x_i$, $E_i(y_i) = y_i$.

Therefore $$x = \sum_{i=1}^{k} E_i(x_i) = \sum_{i=1}^{k} E_i(y_i)$$

By applying E_j to this equation and by using condition b) we get,

$$E_j(x) = E_j^2(x_j) = E_j^2(y_j)$$

$$\therefore \quad x_j = y_j$$

Thus the representation of x is unique.

Therefore ii) $V = W_1 \oplus W_2 \oplus \ldots \oplus W_k$

where $W_i = R_{E_i}$, $i = 1, 2, \ldots, k$

Conversely, let us suppose that $V = W_1 \oplus W_2 \oplus \ldots \oplus W_k$. Then given any $x \in V$, there exists a unique $x_i \in W_i$ such that

$$x = \sum_{i=1}^{k} x_i$$

We define an operator E_i by $E_i(x) = x_i$. Since x_i is unique, E_i is well defined.

Now $$E_i(x+y) = x_i + y_i = E_i(x) + E_i(y)$$

Thus E_i is linear.

Also $$E_i^2(x) = E_i(x_i) = x_i = E_i(x)$$

Thus $$E_i^2 = E_i$$

Thus E_i is a projection. Since $x_i \in W_i$ and $E_i(x_i) = x_i$, E_i is a projection on W_i.

Therefore ii) $W_i = R_{E_i}$, $i = 1, 2, \ldots, k$

To complete the proof we need to show that the projections E_i s satisfy conditions a) and b).

For any $x \in V$, $$x = x_1 + x_2 + \ldots + x_k$$

Since by definition $E_i(x) = x_i$,

$$x = E_1(x) + E_2(x) + \ldots + E_k(x) = (E_1 + E_2 + \ldots + E_k)x$$

Thus $I = E_1 + E_2 + \ldots + E_k$, and condition a) is satisfied.

Suppose $x \in W_j$ and $j \neq i$. Then since $W_j \cap W_i = 0$ and $E_j(x) = x$,

$$E_iE_j(x) = E_i(x) = 0$$

Therefore $E_iE_j = 0$ if $i \neq j$, and condition b) is satisfied ■

THE MINIMUM POLYNOMIAL OF A LINEAR OPERATOR

Let T be a linear operator on V. We shall show that there exists a polynomial in T, say $p(T)$, with $p(T)x = 0$ for every $x \in V$.

Let us assume that $\{x_1, x_2, \ldots, x_n\}$ is a basis for V. Now consider the set of n+1 vectors $\{x_1, T(x_1), T^2(x_1), \ldots, T^n(x_1)\}$. Since V is of dimension n, any set of n+1 vectors will be linearly dependent.

Thus there exist scalars α_{ij}, $j = 0, 1, 2, \ldots, n$, such that the α_{ij} s are not all zero, but

$$\sum_{j=0}^{n} \alpha_{1j} T^j(x_1) = 0$$

Since there exists a largest integer k, such that $1 \le k \le n$, for which $\alpha_{1k} \neq 0$, the above linear combination can be written as

$$\sum_{j=0}^{k} \alpha_{1j} T^j(x_1) = 0$$

By expansion we get

$$\alpha_{10}x_1 + \alpha_{11}T(x_1) + \ldots + \alpha_{1k}T^k(x_1) = 0$$

Then
$$T^k(x_1) = \frac{-1}{\alpha_{1k}} \sum_{j=0}^{k-1} \alpha_{1j} T^j(x_1)$$

or
$$\left(T^k + \frac{1}{\alpha_{1k}} \sum_{j=0}^{k-1} \alpha_{1j} T^j\right) x_1 = 0$$

Now let $$F_1(T) = T^k + \frac{1}{\alpha_{1k}} \sum_{j=0}^{k-1} \alpha_{1j} T^j$$

Then $F_1(T)\, x_1 = 0$. (Notice that $F_1(T)$ is a monic polynomial). Similarly, we can find monic polynomials in T such that

$$F_2(T)\, x_2 = 0$$

$$F_3(T)\, x_3 = 0$$

$$\vdots$$

$$F_n(T)\, x_n = 0$$

Then since $\{x_1, x_2, \ldots, x_n\}$ is a basis for V, we can construct a polynomial in T which annihilates any vector in V as follows:

Let $$p(T) = \prod_{i=1}^{n} F_i(T)$$

Then given any $x \in V$, where in terms of the basis

$$x = \sum_{k=1}^{n} \beta_k x_k$$

we get $$p(T)\, x = \prod_{i=1}^{n} F_i(T) \sum_{k=1}^{n} \beta_k x_k$$

Using the properties of commutativity and linearity of polynomials we can write

$$p(T)\, x = \sum_{k=1}^{n} \alpha_k \Big(\prod_{\substack{i=1 \\ i \neq k}}^{n} F_i(T) \Big) F_k(T)\, x_k$$

Since $F_k(T)\, x_k = 0$, $k = 1, 2, \ldots, n$, $p(T)\, x = 0$ for every $x \in V$ □

Now since there exists a monic polynomial in T, $p(T)$, with $p(T)\, x = 0$ for every $x \in V$, it follows that there exists a monic polynomial of smallest degree, say $\varpi(T)$, such that $\varpi(T)\, x = 0$ for every $x \in V$. This minimal polynomial $\varpi(T) = 0$ is called the **minimum polynomial** of T.

We have just proved the following:

Theorem 6

Every linear operator has a minimum polynomial ■

EIGENVALUES AND EIGENVECTORS

We now focus on a special type of invariant subspace of the operator T on V. When T operates on a vector in this special subspace, the vector gets multiplied by a scalar. We shall prove an important relationship between this scalar and the minimum polynomial of T.

Definition

λ is an **eigenvalue** of T an operator on V if there exists an $x \in V$ with $x \neq 0$, such that $T(x) = \lambda x$.

If λ is an eigenvalue of T and $T(x) = \lambda x$, then x is an **eigenvector** of T corresponding to the eigenvalue λ.

Now given $\varpi(T)$, the minimum polynomial of T, we can construct a polynomial $\varpi(t)$ over the complex field $\mathbb{C}$, which has the same coefficients as those of $\varpi(T)$.

Example: $$\varpi(T) = T^4 + 3T^3 - 7T^2 + T - 4I$$

$$\varpi(t) = t^4 + 3t^3 - 7t^2 + t - 4$$

Since $\mathbb{C}$ is an algebraically closed field, we can write $\varpi(t)$ as the product of linear factors:

$$\varpi(t) = \prod_{j=1}^{k} (t - \lambda_j)^{r_j}$$

where the λ_j s, j = 1, 2, ... , k, are the distinct roots of $\varpi(t) = 0$. We shall show that the λ_j s are the eigenvalues of T.

We define $$\varpi_i(t) = \left(\prod_{\substack{j=1 \\ j \neq i}}^{k} (t - \lambda_j)^{r_j} \right) (t - \lambda_i)^{r_i - 1}$$

Thus $$\varpi(t) = (t - \lambda_i)\, \varpi_i(t)$$

Now since $\varpi_i(t)$ is of degree less than $\varpi(t)$, and given that $\varpi(t)$ corresponds to the minimum polynomial $\varpi(T)$, it follows that there exists at least one vector $x_i \in V$ such that

$$\varpi_i(T)\, x_i \neq 0$$

But since $\varpi(T)$ is the minimum polynomial,

$$\varpi(T)\, x_i = 0$$

Therefore $$(T - \lambda_i I)\,[\varpi_i(T)\,x_i] = 0 \qquad \square$$

Thus we have shown that a root of $\varpi(t)$ is an eigenvalue of T.

Lemma

If λ_1 and λ_2 are distinct eigenvalues of T and x_1 and x_2 are their corresponding eigenvectors, then $x_1 \neq x_2$.

Proof

Suppose x is an eigenvector of T with corresponding eigenvalues λ_1 and λ_2. Then

$$T(x) = \lambda_1 x \quad \text{and} \quad T(x) = \lambda_2 x$$

Thus $$\lambda_1 x = \lambda_2 x$$

Since $x \neq 0$, $$\lambda_1 = \lambda_2$$

The lemma follows by contraposition ■

Theorem 7

λ is a root of $\varpi(t) = 0$, where $\varpi(T)$ is the minimum polynomial of T, if and only if λ is an eigenvalue of T.

Proof

Suppose λ is a root of $\varpi(t) = 0$. We showed above that λ is an eigenvalue of T.

Conversely, suppose λ is an eigenvalue of T. We must show that λ is a root of $\varpi(t) = 0$.

The proof is by contradiction. Let us assume that $\mathfrak{m}(T)$ is the minimum polynomial of T and that λ is an eigenvalue of T such that λ is **not** a root of $\mathfrak{m}(t) = 0$. We show that this is impossible. We can factor $\mathfrak{m}(t)$ into linear factors that are not necessarily distinct as follows:

$$\mathfrak{m}(t) = \prod_{j=1}^{r} (t - \lambda_j)$$

Let $$\mathfrak{m}_k(t) = \prod_{j=1}^{k} (t - \lambda_j) \qquad \text{where } k \le r - 1$$

Now let x be an eigenvector of T corresponding to the eigenvalue λ. Since λ is not a zero of $\mathfrak{m}(t) = 0$, $\lambda \ne \lambda_j$ for every $j = 1, 2, \ldots, r$.

Then either

i) $\mathfrak{m}_k(T)x = 0$ for all $k = 1, 2, \ldots, r - 1$

or ii) $\mathfrak{m}_k(T)x \ne 0$ for some $k = 1, 2, \ldots, r - 1$

Suppose i) $\mathfrak{m}_k(T)x = 0$ for all $k = 1, 2, \ldots, r - 1$.

Then in particular,

$$(T - \lambda_1 I)x = 0$$

Thus x is an eigenvector of T corresponding to the distinct eigenvalues λ_1 and λ. But by the lemma, this is impossible.

On the other hand, suppose ii) $\mathfrak{m}_k(T)x \ne 0$ for some $k = 1, 2, \ldots, r - 1$. Then

there is a largest $k \le r - 1$ such that $\mathfrak{m}_k(T)x \neq 0$ but $\mathfrak{m}_{k+1}(T)x = 0$.

Then $$\mathfrak{m}_{k+1}(T)x = (T - \lambda_{k+1}I)\,[\mathfrak{m}_k(T)x] = 0$$

This implies that $\mathfrak{m}_k(T)x$ is an eigenvector of T corresponding to the eigenvalue λ_{k+1}. But since x is an eigenvector of T corresponding to the eigenvalue λ,

$$(T - \lambda I)\,[\mathfrak{m}_k(T)x] = [\mathfrak{m}_k(T)]\,(T - \lambda I)\,x = \mathfrak{m}_k(T).\,0$$

Therefore $$(T - \lambda I)\,[\mathfrak{m}_k(T)x] = 0$$

Thus $\mathfrak{m}_k(T)x$ is an eigenvector of T corresponding to the distinct eigenvalues λ_{k+1} and λ, which again is impossible ■

Example

Let the matrix representation of the operator T on $\mathbb{C}^4$ with respect to the basis $\{x_1, x_2, x_3, x_4\}$ be given by

$$[\,T\,] = \begin{bmatrix} 2 & 1 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & -1 & -1 & 0 \\ 0 & 0 & 0 & 2 \end{bmatrix}$$

Then $$T(x_1) = 2\,x_1$$

and $$T(x_1) - 2\,x_1 = 0$$

Thus $$F_1(t) = t - 2$$

Similarly,

$$T(x_3) = x_1 - x_3$$

$$T(x_3) + x_3 = x_1$$

$$T^2(x_3) + T(x_3) = T(x_1) = 2x_1 = 2[T(x_3) + x_3]$$

$$T^2(x_3) - T(x_3) - 2x_3 = 0$$

Thus

$$F_3(t) = t^2 - t - 2$$

$$F_3(t) = (t-2)(t+1)$$

Also

$$T(x_2) = x_1 + x_2 - x_3$$

$$T(x_2) - x_2 = x_1 - x_3 = T(x_3)$$

$$T^2(x_2) - T(x_2) = T^2(x_3)$$

$$T^2(x_2) - T(x_2) - [T(x_2) - x_2] = T^2(x_3) - T(x_3)$$

$$T^2(x_2) - 2T(x_2) + x_2 = T^2(x_3) - T(x_3) = 2x_3$$

$$T^3(x_2) - 2T^2(x_2) + T(x_2) = 2T(x_3) = 2[T(x_2) - x_2]$$

$$T^3(x_2) - 2T^2(x_2) - T(x_2) + 2x_2 = 0$$

Thus

$$F_2(t) = t^3 - 2t^2 - t + 2$$

$$F_2(t) = (t-2)(t-1)(t+1)$$

Finally

$$T(x_4) = 2x_4$$

$$T(x_4) - 2x_4 = 0$$

Thus

$$F_4(t) = t - 2$$

Now $$p(t) = \prod_{i=1}^{4} F_i(t) = (t-2)^4 (t-1)(t+1)^2$$

Thus $$p(T) = (T-2I)^4 (T-I)(T+I)^2$$

Observe that

$$F_2(T)\, x_i = (T-2I)(T-I)(T+I)\, x_i = 0 \qquad i = 1, 2, 3, 4$$

Then the minimum polynomial of T, $\mathfrak{m}(T)$, is in this case $F_2(T)$.

Thus $$\mathfrak{m}(T) = (T-2I)(T-I)(T+I)$$

Therefore the eigenvalues of T are 2, 1, and -1.

N.B. The degree of $\mathfrak{m}(T) \leq$ the dimension of V.

THE CHARACTERISTIC POLYNOMIAL OF A LINEAR OPERATOR

In the next section, we shall present the main theorem used in constructing the Jordan form, a theorem which depends on the idea of the minimum polynomial. But first we elaborate on the observation that the degree of $\mathfrak{m}(T) \leq$ the dimension of V.

Observe that in our example above,

$$\mathfrak{m}(T) = (T-2I)(T-I)(T+I)$$

Thus the degree of $\mathfrak{m}(T)$ is 3, while V is of dimension 4.

As implied in the Introduction, the special basis we shall construct for the Jordan form consists of certain invariant basis elements, namely, the eigenvectors of $\mathcal{T}$. As basis elements, these eigenvectors must be linearly independent and must span V.

However, although $\boldsymbol{\varpi}(\mathcal{T})$ always indicates the **distinct** eigenvalues of $\mathcal{T}$ (it may also repeat some), as with our example, the **number** of roots of $\boldsymbol{\varpi}(t)$ need not equal the dimension of V. Thus given the minimum polynomial of a linear operator, we cannot always determine which eigenvalue(s) is (are) repeated. Therefore we may be unable to expand our initial number of eigenvectors (= number of roots of $\boldsymbol{\varpi}(t)$) to an eigenbasis which spans V.

It is the characteristic polynomial of $\mathcal{T}$ which always provides us with the means of determining which eigenvalues are repeated and the number of such repetitions.

Definition

Let $[\,T\,]$ be a matrix representation of the linear operator $\mathcal{T}$ on the n-dimensional vector space V. If

$$\mathbf{c}(\lambda) = \mathbf{det}\,(\,[\,T\,] - \lambda\, I\,)$$

then $\mathbf{c}(\lambda)$ is called the **characteristic polynomial** of $\mathcal{T}$.

We compute $\mathbf{c}(\lambda)$ for our previous example:

$$\det(\,[\,T\,] - \lambda\, I\,) = \det \begin{bmatrix} 2-\lambda & 1 & 1 & 0 \\ 0 & 1-\lambda & 0 & 0 \\ 0 & -1 & -1-\lambda & 0 \\ 0 & 0 & 0 & 2-\lambda \end{bmatrix}$$

$$c(\lambda) = (2 - \lambda)(1 - \lambda)(-1 - \lambda)(2 - \lambda)$$

Our discussion of the characteristic polynomial of a linear operator is limited to observing that the roots of this polynomial provide dim(V) eigenvalues (not necessarily distinct) of T. In our example, the roots of $c(\lambda)$ are 2, 1, -1, and 2. Thus the eigenvalue "missing" from the roots of $\omega(T)$ is a repeated 2. The characteristic polynomial is thus a way to find "all" the eigenvalues and eigenvectors of a linear operator. However, finding the determinant of a n x n matrix is not practical for values of n exceeding 3 or 4, since it requires $n!(n^2 + n)$ multiplications.

Therefore, we state, but do not prove the following theorem:

Theorem 8

Let T be a linear operator on a vector space V. Then the roots of $c(\lambda) = 0$ are the eigenvalues of T ■

N. B. Eigenvalues and eigenvectors are sometimes called **characteristic values** and **characteristic vectors**, since "eigen-" is the German prefix which means "characteristic" (belonging to). The characteristic polynomial of a linear operator is the embodiment of these invariant characteristics.

THE DECOMPOSITION OF AN OPERATOR

We now present the main theorem used in constructing the Jordan form.

Theorem 9

Let $\varpi(T)$ be the minimum polynomial of T on the vector space V over the algebraically closed field $\mathbb{F}$. If

$$\varpi(t) = \prod_{j=1}^{k} (t - \lambda_i)^{r_i}$$

where $\lambda_i \neq \lambda_j$ for $i \neq j$ and $r_i > 0$ for $i = 1, 2, \ldots, k$, then

i) $V = U_1 \oplus U_2 \oplus \ldots \oplus U_k$ where U_i is the null space of $(T - \lambda_i I)^{r_i}$

ii) U_i is invariant under T for each i.

iii) If T_i is the restriction of T to U_i, then $(T - \lambda_i I)^{r_i}$ is the minimum polynomial of T_i.

Proof

Let us define $\varpi_i(t)$ by

$$\varpi_i(t) = \frac{\varpi(t)}{(t - \lambda_i)^{r_i}} = \prod_{\substack{j=1 \\ j \neq i}}^{k} (t - \lambda_j)^{r_j}$$

Then the greatest common devisor of $\{ \varpi_1(t), \varpi_2(t), \ldots, \varpi_k(t) \}$ is 1.

Thus there exist polynomials $q_i(t)$ such that

$$\sum_{i=1}^{k} q_i(t)\, \varpi_i(t) = 1$$

We write this identity in t as an identity in T.

Thus
$$\sum_{i=1}^{k} q_i(T)\, \varpi_i(T) = I$$

If we let $E_i = q_i(T)\, \varpi_i(T)$, then

Thus
$$\sum_{i=1}^{k} E_i = I$$

Notice that

$$\varpi_i(T)\, \varpi_j(T) = \frac{\varpi(T)}{(T - \lambda_i I)^{r_i}} \cdot \frac{\varpi(T)}{(T - \lambda_j I)^{r_j}}$$

Therefore for $i \neq j$,
$$\varpi_i(T)\, \varpi_j(T) = \varpi(T) \prod_{\substack{j=1 \\ j \neq i}}^{k} (T - \lambda_p I)^{r_p}$$

Hence, if $i \neq j$, $q_i(T)\, \varpi_i(T)\, q_j(T)\, \varpi_j(T) = 0$, since it is a multiple of $\varpi(T)$. In other words,

$$E_i E_j = 0 \qquad \text{if } i \neq j$$

Also we have shown that

$$E_1 + E_2 + \ldots + E_k = I$$

By Theorem 5, these conditions imply that the E_i s are projections on V which lead to the following direct sum decomposition of V:

$$V = U_1 \oplus U_2 \oplus \dots \oplus U_k \qquad \text{where } U_i = R_{E_i}$$

To complete the proof we must prove that the U_i s fulfill the three conditions described in the theorem.

i) We must first show that U_i is the null space of $(T - \lambda_i I)^{r_i}$, $i = 1, 2, \dots, k$.

Let $x \in U_i$. Then $$E_i(x) = x$$

Thus $$(T - \lambda_i I)^{r_i} x = (T - \lambda_i I)^{r_i} E_i(x)$$

$$(T - \lambda_i I)^{r_i} x = \frac{\varpi(T)}{\varpi_i(T)} q_i(T)\, \varpi_i(T)\, x = \varpi(T)\, q_i(T)\, x = q_i(T)\, \varpi(T)\, x$$

Then $$(T - \lambda_i I)^{r_i} x = 0$$

Thus x belongs to the null space of $(T - \lambda_i I)^{r_i}$. Therefore $(T - \lambda_i I)^{r_i} \supset U_i$.

On the other hand, suppose x is in the null space of $(T - \lambda_i I)^{r_i}$.

Since $$\varpi_j(t) = \prod_{\substack{i=1 \\ i \neq j}}^{k} (T - \lambda_i I)^{r_i}$$

for $j \neq i$, $q_j(T)\, \varpi_j(T)$ is a multiple of $(T - \lambda_i I)^{r_i}$.

Hence $E_j(x) = q_j(T)\, \varpi_j(T)\, x = 0$ since x is in the null space of $(T - \lambda_i I)^{r_i}$.

But if $E_j(x) = 0$, for $i \neq j$ $\qquad E_i(x) = x$

Then $\qquad x \in U_i$

Thus $U_i \supset$ the null space of $(T - \lambda_i I)^{r_i}$. But the null space of $(T - \lambda_i I)^{r_i} \supset U_i$.

Therefore U_i is the null space of $(T - \lambda_i I)^{r_i}$, $i = 1, 2, \ldots, k$.

ii) Next we must show that U_i is invariant under T for $i = 1, 2, \ldots, k$.

i.e. we must show that if $x \in U_i$, then $T(x) \in U_i$. Note that since E_i is a polynomial in T, T commutes with E_i.

Let $x \in U_i$. Then $\qquad E_i(x) = x$

Thus $\qquad E_i(x) \in U_i$

Now $\qquad T E_i(x) = T(x)$

then $\qquad E_i [T(x)] = T(x)$

But $E_i [T(x)] \in U_i$. Thus $\qquad T(x) \in U_i$

Therefore each U_i is invariant under T, $i = 1, 2, \ldots, k$.

iii) Finally, we must show that if T_i is the restriction of T to U_i, then $(T - \lambda_i I)^{r_i}$ is the minimum polynomial of T_i.

Let T_i be the restriction of T to U_i. Then by condition i) since U_i is the null space of $(T - \lambda_i I)^{r_i}$, $\qquad (T - \lambda_i I)^{r_i} x = 0 \qquad$ for all $x \in U_i$

Let $s(T)$ be the minimum polynomial of T_i. Then $s(T)$ divides $(T - \lambda_i I)^{r_i}$. (We

note that in terms of T_i, $s(T_i)$, the minimum polynomial of T_i , divides

$(T_i - \lambda_i I)^{r_i}$).

Then for any $x \in U_i$, $\qquad s(T)x = 0$

Also for any $x \in U_1 \oplus U_2 \oplus \ldots \oplus U_{i-1} \oplus U_{i+1} \oplus \ldots \oplus U_k$,

$$\omega_i(T)x = 0$$

Thus for $x \in V$, $\qquad s(T)\,\omega_i(T)x = 0$

But since $\omega(T)$ is the minimum polynomial of T, it follows that $\omega(T)$ divides $s(T)\,\omega_i(T)$. However, by definition $\omega(T) = (T - \lambda_i I)^{r_i}\,\omega_i(T)$. Thus $(T - \lambda_i I)^{r_i}\,\omega_i(T)$ divides $s(T)\,\omega_i(T)$.

In other words, $\qquad (T - \lambda_i I)^{r_i}$ divides $\omega_i(T)$

But by our original assumption, $s(T)$ divides $(T - \lambda_i I)^{r_i}$.

Therefore $\qquad s(T) = (T - \lambda_i I)^{r_i}$

where $s(T)$ is the minimum polynomial of T_i, the restriction of T to U_i ■

We observe that the theorem implies that U_i is a summand of V associated with the eigenvalue λ_i, if and only if $(T - \lambda_i I)$ is nilpotent on U_i. Thus when T is restricted to the subspace U_i, it has only one eigenvalue, namely λ_i. The subspace U_i is called an **eigenspace** of V.

THE SPECIAL BASIS FOR THE JORDAN FORM OF A NILPOTENT OPERATOR

In this section we shall construct the special basis for a vector space which enables us to determine the Jordan matrix representation of a linear operator that is nilpotent on that vector space. As we shall show, the construction will be analogous to the construction of the special basis which gives a Jordan representation of any linear operator.

First we examine the decomposition of a vector space in greater detail.

Suppose T is an operator on V that is nilpotent of index k on V. Then $T^k = 0$. Let N_j be the null space of T^j. We shall show that for $i < j$, $N_j \supset T^i(N_j)$. In other words, we shall show that N_j is invariant under T^i if $i < j$.

Let $y \in N_j$. We must show that $T^j(T^i(y)) = 0$.

$$T^j(T^i(y)) = T^i(T^j(y)) = T^i(0) \qquad \text{since } y \in N_j$$

and

$$T^i(0) = 0$$

Therefore N_j is invariant under T^i if $i < j$. In particular, N_j is invariant under T^{j-1}.

Observe that

$$N_j \supset N_{j-1}$$

Let $\{x_i\}_{i=1}^{r}$ be a basis for N_{j-1}. Complete this to a basis for N_j, say $\{x_i\}_{i=1}^{k}$.

Define P_{j-1} to be the subspace of N_j, spanned by the basis $\{x_i\}_{i=r+1}^{k}$.

Then

$$N_j = N_{j-1} \oplus P_{j-1}$$

Observe that if $x \in P_{j-1}$ and $x \neq 0$, then

$$T^j(x) = 0 \text{ but } T^{j-1}(x) \neq 0$$

Now since T is nilpotent of index k on V, it follows that

$$\begin{aligned} V = N_k &= P_{k-1} \oplus N_{k-1} \\ &= P_{k-1} \oplus P_{k-2} \oplus N_{k-2} \end{aligned}$$

Then by induction,

$$V = P_{k-1} \oplus P_{k-2} \oplus \ldots \oplus P_1 \oplus P_0 \oplus N_0$$

But $N_0 = \{0\}$

Therefore $\quad V = P_{k-1} \oplus P_{k-2} \oplus \ldots \oplus P_1 \oplus P_0 \quad \square$

The following lemma helps determine the special basis for V in terms of the bases for P_i, $i = 1, 2, \ldots, k-1$.

Lemma

If $\{x_i\}_{i=1}^{p}$ is a basis for P_j, then a basis for P_{j-1} can be chosen to include $\{T(x_i)\}_{i=1}^{p}$.

Proof

Let $\{x_i\}_{i=1}^{p}$ be a basis for P_j and $\{y_i\}_{i=1}^{q}$ be a basis for N_{j-1}. We first show that $\mathcal{Y} = \{ y_1, y_2, \dots, y_q, T(x_1), T(x_2), \dots, T(x_p) \}$ is linearly independent.

Suppose $\mathcal{Y}$ is linearly dependent. Then we can have

$$\sum_{i=1}^{q} \alpha_i y_i + \sum_{i=1}^{p} \beta_i T(x_i) = 0$$

where the α_i s and β_i s are not all zero.

Since $\{y_i\}_{i=1}^{q}$ is linearly independent being a basis for N_{j-1}, $\alpha_i \neq 0$ for $i = 1, 2, \dots, q$, unless $\sum_{i=1}^{q} \alpha_i y_i = 0$.

For β_i, $i = 1, 2, \dots, p$, there exists an integer k such that $0 < k \leq p$ and $\beta_k \neq 0$.

Thus
$$\beta_k T(x_k) = \sum_{i=1}^{q} \alpha_i y_i + \sum_{i=1}^{k-1} \beta_i T(x_i)$$

We operate on this equation by T^{j-1}. Then since $\{y_i\}_{i=1}^{q}$ is in N_{j-1}, we get

$$\beta_k T^j(x_k) = \sum_{i=1}^{k-1} \beta_i T^j(x_i)$$

Let $\gamma_i = (\beta_i / \beta_k)$.

Then
$$T^j(x_k - \sum_{i=1}^{k-1} \gamma_i x_i) = 0$$

Thus $$x_k - \sum_{i=1}^{k-1} \gamma_i x_i \in N_j$$

Since $\{x_i\}_{i=1}^{p}$ is a basis for P_j, and $0 < k \leq p$,

$$x_k - \sum_{i=1}^{k-1} \gamma_i x_i \in P_j$$

Then $$x_k - \sum_{i=1}^{k-1} \gamma_i x_i \in P_j \cap N_j$$

But $P_j \cap N_j = \emptyset$.

Hence $$x_k - \sum_{i=1}^{k-1} \gamma_i x_i = 0$$

But this implies that $\{x_i\}_{i=1}^{k}$ is a linearly dependent set, which is impossible since for $p \geq k$, $\{x_i\}_{i=1}^{p}$ is a linearly independent set, because it is a basis for P_j.

By this contradiction, it follows that

$$\mathcal{Y} = \{ y_1, y_2, \dots, y_q, T(x_1), T(x_2), \dots, T(x_p) \} \text{ is linearly independent.}$$

Now since $T^j[T(x_i)] = 0$, $i = 1, 2, \dots, p$, and $T^j(y_k) = 0$, $k = 1, 2, \dots, q$, $\mathcal{Y}$ is a subspace of N_j. We complete $\mathcal{Y}$ to a basis for N_j, say

$$\mathcal{Y}_1 = \{ y_1, y_2, \dots, y_q, T(x_1), T(x_2), \dots, T(x_p), z_1, z_2, \dots, z_m \}$$

We know that $\{ y_1, y_2, \dots, y_q \}$ is a basis for N_{j-1}.

Then since $$N_j = N_{j-1} \oplus P_{j-1}$$

it follows that $\{ T(x_1), T(x_2), \ldots, T(x_p), z_1, z_2, \ldots, z_m \}$ is a basis for P_{j-1} ■

We now give an example to illustrate the special basis for the Jordan form of a nilpotent operator on V. Following the example we shall give a detailed construction of this special basis.

Example

Let T be an operator on a vector space V that has the decomposition

$$V = P_3 \oplus P_2 \oplus P_1 \oplus P_0$$

where if $x \in P_{j-1}$ and $x \neq 0$, then

$$T^j(x) = 0 \text{ but } T^{j-1}(x) \neq 0$$

Also it follows that T is nilpotent on V of index 4.

Suppose that P_3 is of dimension 2, P_2 is of dimension 3, P_1 is of dimension 5, and P_0 is of dimension 7.

Let $\{P_3\} = \{ x_1, x_2 \}$ be a basis for P_3. Then by the lemma, it follows that

$\{P_2\} = \{ T(x_1), T(x_2), x_3 \}$ will be a basis for P_2,

$\{P_1\} = \{ T^2(x_1), T^2(x_2), T(x_3), x_4, x_5 \}$ will be a basis for P_1, and

$\{P_0\} = \{ T^3(x_1), T^3(x_2), T^2(x_3), T(x_4), T(x_5), x_6, x_7 \}$ will be a basis for P_0.

By the decomposition of V it follows that $\{P_3\} \cup \{P_2\} \cup \{P_1\} \cup \{P_0\}$ is a basis for V. We write this basis in the following row and column format:

$\{P_3\}$	$\{P_2\}$	$\{P_1\}$	$\{P_0\}$
x_1	$T(x_1)$	$T^2(x_1)$	$T^3(x_1)$
x_2	$T(x_2)$	$T^2(x_2)$	$T^3(x_2)$
	x_3	$T(x_3)$	$T^2(x_3)$
		x_4	$T(x_4)$
		x_5	$T(x_5)$
			x_6
			x_7

Let us consider the operation of T on the basis elements of V in the following **special order** : read the elements of the basis from the top across each row from right to left.

Then by the definition of P_j and since $T^4 = 0$, the matrix representation of T with respect to the basis elements from the first row is

$$[\,Q\,] = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

Similarly, the matrix representation of T with respect to the basis elements from the second row is identical to $[\,Q\,]$.

In a similar manner, the matrix representation of T with respect to the basis elements from the third row is

$$[\,R\,] = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}$$

The matrix representations of $\mathcal{T}$ with respect to the basis elements from the fourth and fifth rows are each of the form

$$[S] = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$$

And finally, the matrix representations of $\mathcal{T}$ with respect to the basis elements from the sixth and seventh rows are each the 1 x 1 zero matrix.

We now combine the matrices to obtain the following Jordan matrix representation of our linear operator $\mathcal{T}$ with respect to the special basis obtained above.

$$[T] = \begin{bmatrix}
0 & 1 & 0 & 0 & & & & & & & & & & & \\
0 & 0 & 1 & 0 & & & & & & & & & & & \\
0 & 0 & 0 & 1 & & & & & & & & & & & \\
0 & 0 & 0 & 0 & & & & & & & & & & & \\
 & & & & 0 & 1 & 0 & 0 & & & & & & & \\
 & & & & 0 & 0 & 1 & 0 & & & & & & & \\
 & & & & 0 & 0 & 0 & 1 & & & & & & & \\
 & & & & 0 & 0 & 0 & 0 & & & & & & & \\
 & & & & & & & & 0 & 1 & 0 & & & & \\
 & & & & & & & & 0 & 0 & 1 & & & & \\
 & & & & & & & & 0 & 0 & 0 & & & & \\
 & & & & & & & & & & & 0 & 1 & & \\
 & & & & & & & & & & & 0 & 0 & & \\
 & & & & & & & & & & & & & 0 & 1 \\
 & & & & & & & & & & & & & 0 & 0 \\
 & & & & & & & & & & & & & & & 0 \\
 & & & & & & & & & & & & & & & & 0
\end{bmatrix}$$

Our example gives us a visual idea of the simple Jordan representation of a linear operator **which is nilpotent** on the vector space. However, our object is to give a simple representation for **any** linear operator on a vector space. In order to do so we first describe the construction of the special basis illustrated in the example above for an arbitrary nilpotent operator V. We shall then construct the Jordan matrix of any linear operator $\mathcal{T}$ with respect to the implication of Theorem 9, which was that U_i is an eigenspace of V associated with the eigenvalue λ_i, if and only if $(\mathcal{T} - \lambda_i I)$ is nilpotent on U_i.

Thus we return to the decomposition of V into

$$V = P_{k-1} \oplus P_{k-2} \oplus \ldots \oplus P_1 \oplus P_0$$

and use the lemma to construct the basis for the Jordan form as follows:

Let $\{x_{i,k-1}\}_{i=1}^{p}$ be a basis for P_{k-1} such that

$$x_{i,k-2} = \mathcal{T}(x_{i,k-1})$$

Then by the lemma, we can complete the following basis for P_{k-2}.

$$\{P_{k-2}\} = \left\{ \{\mathcal{T}(x_{i,k-1})_{i=1}^{p}\}, \{x_{i,k-2}\}_{i=p+1}^{p+q} \right\}$$

We proceed by induction until a basis for P_0 has been obtained.

Then since

$$V = P_{k-1} \oplus P_{k-2} \oplus \ldots \oplus P_1 \oplus P_0$$

it follows that the set consisting of the bases for P_j, $j = 0, 1, 2, \ldots, k-1$, is a basis for V. Let $\{P_j\}$ be the basis for P_j. Then we can represent the basis for V under the columns $\{P_j\}$, $j = 0, 1, 2, \ldots, k-1$, as follows:

$\{P_{k-1}\}$	$\{P_{k-2}\}$	$\{P_{k-3}\}$		$\{P_1\}$	$\{P_0\}$
$x_{1,k-1}$	$\pi(x_{1,k-1})$	$\pi^2(x_{1,k-1})$	.....	$\pi^{k-2}(x_{1,k-1})$	$\pi^{k-1}(x_{1,k-1})$
$x_{2,k-1}$	$\pi(x_{2,k-1})$	$\pi^2(x_{2,k-1})$	.....	$\pi^{k-2}(x_{2,k-1})$	$\pi^{k-1}(x_{2,k-1})$
$\vdots$	$\vdots$	$\vdots$		$\vdots$	$\vdots$
$\vdots$	$\vdots$	$\vdots$		$\vdots$	$\vdots$
$x_{p,k-1}$	$\pi(x_{p,k-1})$	$\pi^2(x_{p,k-1})$	.....	$\pi^{k-2}(x_{p,k-1})$	$\pi^{k-1}(x_{p,k-1})$
	$x_{p+1,k-2}$	$\pi(x_{p+1,k-2})$	.....	$\pi^{k-3}(x_{p+1,k-2})$	$\pi^{k-2}(x_{p+1,k-2})$
	$\vdots$	$\vdots$		$\vdots$	$\vdots$
	$\vdots$	$\vdots$		$\vdots$	$\vdots$
	$x_{p+q,k-2}$	$\pi(x_{p+q,k-2})$	.....	$\pi^{k-3}(x_{p+q,k-2})$	$\pi^{k-2}(x_{p+q,k-2})$
		$x_{p+q+1,k-3}$	.....	$\pi^{k-4}(x_{p+q+1,k-3})$	$\pi^{k-3}(x_{p+q+1,k-3})$
		$\vdots$		$\vdots$	$\vdots$
		$\vdots$		$\vdots$	$\vdots$
		$x_{p+q+r,k-3}$	.....	$\pi^{k-4}(x_{p+q+r,k-3})$	$\pi^{k-3}(x_{p+q+r,k-3})$
			.....	$\vdots$	$\vdots$
			.....	$\vdots$	$\vdots$
			.....	$\vdots$	$\vdots$
				$\vdots$	$\vdots$
				$x_{p+q+r+\ldots+m,1}$	$\pi(x_{p+q+r+\ldots+m,1})$
					$x_{p+q+r+\ldots+m+1,0}$
					$\vdots$
					$\vdots$
					$x_{p+q+r+\ldots+m+n,0}$

The Jordan form of a matrix results from a special ordering of the above basis elements of V.

Consider the basis for V as consisting of the sets of t-elements,

t = k, k − 1, ... , 2, 1, obtained from the bases $\{P_0\}$, $\{P_1\}$, ... , $\{P_{k-2}\}$, $\{P_{k-1}\}$ as follows:

Each set will have a **t** number of basis elements, **t** = k, k − 1, ... , 2, 1, which are obtained by reading the entries in the columns $\{P_j\}$, j = 0, 1, 2, ... , k − 1, across each row from right to left.

For example, the first **p** rows will yield a **p** number of sets with each set containing a **k** number of basis elements of the form

$$\mathcal{X}_i = \{ \mathcal{T}^{k-1}(x_{i,k-1}), \mathcal{T}^{k-2}(x_{i,k-1}), \ldots , \mathcal{T}^2(x_{i,k-1}), \mathcal{T}(x_{i,k-1}), x_{i,k-1}) \}$$

where each i, i = 1, 2, ... , p, corresponds to the specific row. Let U_i be the subspace covered by the complete basis $\{ \mathcal{X}_1, \mathcal{X}_2, \ldots , \mathcal{X}_p \}$.

Notice that $\mathcal{T}$ operating on the basis $\mathcal{X}_i$, annihilates $\mathcal{T}^{k-1}(x_{i,k-1})$, the element from the basis $\{P_o\}$, but $\mathcal{T}$ operating on $\mathcal{T}^{k-2}(x_{i,k-1})$, the element from the basis $\{P_1\}$, replaces the element from $\{P_o\}$.

Similarly, the element from $\{P_2\}$ replaces that from $\{P_1\}$ when operated on by $\mathcal{T}$. In fact, we can represent the operation of $\mathcal{T}$ on any particular set of basis elements $\mathcal{X}_i$, by the following matrix :

$$
[Q_i] = \begin{array}{c|ccccccc}
\mathcal{T} & \mathcal{T}^{k-1}(x_{i,k-1}) & \mathcal{T}^{k-2}(x_{i,k-1}) & \mathcal{T}^{k-3}(x_{i,k-1}) & \cdots\cdots & \mathcal{T}^{2}(x_{i,k-1}) & \mathcal{T}(x_{i,k-1}) & x_{i,k-1} \\
\hline
\mathcal{T}^{k-1}(x_{i,k-1}) & 0 & 1 & 0 & \cdots\cdots & 0 & 0 & 0 \\
\mathcal{T}^{k-2}(x_{i,k-1}) & 0 & 0 & 1 & \cdots\cdots & 0 & 0 & 0 \\
\mathcal{T}^{k-3}(x_{i,k-1}) & 0 & 0 & 0 & & 0 & 0 & 0 \\
\vdots & \vdots & \vdots & \vdots & & \vdots & \vdots & \vdots \\
\vdots & & \vdots & \vdots & & & \vdots & \vdots \\
\vdots & & & \vdots & & & & \vdots \\
\mathcal{T}^{2}(x_{i,k-1}) & 0 & 0 & 0 & \cdots\cdots & 0 & 1 & 0 \\
\mathcal{T}(x_{i,k-1}) & 0 & 0 & 0 & \cdots\cdots & 0 & 0 & 1 \\
x_{i,k-1} & 0 & 0 & 0 & \cdots\cdots & 0 & 0 & 0
\end{array}
$$

$i = 1, 2, \ldots, p$. Observe that $[Q_i]$ is a $k \times k$ matrix.

Let

$$
[T_1] = \begin{bmatrix}
[Q_1] & 0 & 0 & 0 \\
0 & [Q_2] & 0 & 0 \\
0 & 0 & [Q_3] & 0 \\
\vdots & \vdots & \vdots & \vdots \\
\vdots & \vdots & \vdots & \vdots \\
0 & 0 & 0 & [Q_p]
\end{bmatrix}
$$

where each $[Q_i]$, $i = 1, 2, \ldots, p$, is the $k \times k$ matrix obtained above.

Let $\mathcal{T}_1$ be the restriction of $\mathcal{T}$ to U_1. Then $[T_1]$ is the matrix representation of $\mathcal{T}_1$ with respect to the basis $\{\mathcal{X}_1, \mathcal{X}_2, \ldots, \mathcal{X}_p\}$. It is also the case that $\mathcal{T}_1$ is nilpotent on U_1 of index k.

Thus far, $\mathcal{T}$ operating on the basis elements from the first **p** rows of the columns $\{P_j\}$, $j = 0, 1, 2, \ldots, k-1$, has yielded the matrix $[T_1]$.

In a similar manner, we can construct a matrix $[T_2]$, which represents the operation of $\mathcal{T}$ on the next **q** rows of the columns $\{P_j\}$. Since $\{P_{k-1}\}$ has no more entries, $j = 0, 1, 2, \ldots, k-2$. Thus each matrix $[Q_i]$, $i = 1, 2, \ldots, q$, from which we construct $[T_2]$, is a $(k-1) \times (k-1)$ matrix. Also if U_2 is the space spanned by the basis for $[T_2]$, and $\mathcal{T}_2$ is the restriction of $\mathcal{T}_2$ to U_2, then $\mathcal{T}_2$ is nilpotent of index $k-1$ on U_2.

Continuing in this manner, we will arrive at the following matrix representation of $\mathcal{T}$.

$$[T] = \begin{bmatrix} [T_1] & 0 & 0 & 0 \\ 0 & [T_2] & 0 & 0 \\ 0 & 0 & [T_3] & 0 \\ \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & [T_k] \end{bmatrix}$$

$[T]$ will have zero entries everywhere except along the upper off diagonal,

where the entries will be either zero or one.

Observe that $[T_k]$ will consist of the zero matrix, since each matrix $[Q_i]$, $i = 1, 2, \ldots, n$, will be the 1×1 zero matrix :

$$[Q_i] = \begin{array}{c|c} \mathcal{T} & x_{i,0} \\ \hline x_{i,0} & 0 \end{array}$$

Thus $[T_k]$ will be the $n \times n$ zero matrix.

THE JORDAN FORM

When constructing the special basis for a Jordan matrix representation of a linear operator $\mathcal{T}$ on a vector space V, we assumed that $\mathcal{T}$ was nilpotent of index k on V. We now describe how to obtain a Jordan matrix representation for **any** linear operator $\mathcal{T}$ on V. (Note that since V is a complex finite-dimensional vector space, our field is algebraically closed).

Let $V = U_1 \oplus U_2 \oplus \ldots \oplus U_k$, where U_i, $i = 1, 2, \ldots, k$, are the eigenspaces of a vector space V under any linear operator $\mathcal{T}$ on V, and let $\mathcal{T}_i$ be the restriction of $\mathcal{T}$ to U_i. Then from Theorem 9, we know that if λ_i is the eigenvalue associated with U_i, $(\mathcal{T}_i - \lambda_i I)$ is nilpotent of index r_i on U_i.

Thus we can construct a special basis for the eigenspace U_i, such

that the matrix representation of $(\mathcal{T}_i - \lambda_i I)$ on U_i is in the Jordan form for a nilpotent operator, $i = 1, 2, \ldots, k$.

In other words, we consider the following analogies between our present construction of a special basis for V, where $\mathcal{T}$ is any operator on V, and our former construction of a special basis for V where the operator $\mathcal{T}$ was nilpotent of index k on V. We treat

i) each eigenspace U_i as analogous to the vector space V.

ii) each operator $(\mathcal{T}_i - \lambda_i I)$ as analogous to the operator $\mathcal{T}$.

iii) each r_i as analogous to k.

We examine the operation of $(\mathcal{T}_i - \lambda_i I)$ on U_i, for $i = 1, 2, \ldots, k$.

Let $[T_i - \lambda_i I]$ be the Jordan matrix representation of $(\mathcal{T}_i - \lambda_i I)$ on U_i with respect to the special basis constructed for U_i (the construction is similar to our earlier construction with the analogous relationships decribed above). Then it follows that

$$[T_i - \lambda_i I] = \begin{bmatrix} [T_{i1}] & 0 & 0 & 0 \\ 0 & [T_{i2}] & 0 & 0 \\ 0 & 0 & [T_{i3}] & 0 \\ \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & [T_{im}] \end{bmatrix}$$

where
$$[T_{ir}] = \begin{bmatrix} [Q_{r1}] & 0 & 0 & 0 \\ 0 & [Q_{r2}] & 0 & 0 \\ 0 & 0 & [Q_{r3}] & 0 \\ \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & [Q_{rs}] \end{bmatrix}$$

The number s indicates the number of $[Q_{rt}]$'s in each $[T_{ir}]$, $r = 1, 2, \ldots, m$. ($[T_{ir}]$ corresponds to $[T_i]$, $i = 1, 2, \ldots, j$, in our former construction). The $[Q_{rt}]$'s will be of the following form:

$$[Q_{rt}] = \begin{bmatrix} 0 & 1 & 0 & 0 & \cdots & 0 \\ 0 & 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & & \vdots \\ \vdots & \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & 0 & \cdots & 1 \\ 0 & 0 & 0 & 0 & \cdots & 0 \end{bmatrix}$$

The size of each $[Q_{rt}]$, $t = 1, 2, \ldots, s$, is the same for a given $[T_{ir}]$. ($[Q_{rt}]$ corresponds to $[Q_i]$ in our former construction, where $[Q_i]$, $i = 1, 2, \ldots, p$, was a $k \times k$ matrix for $[T_1]$).

We use the following notation :
$$\bigoplus_r [T_{ir}] = \begin{bmatrix} [T_{i1}] & 0 & 0 & 0 \\ 0 & [T_{i2}] & 0 & 0 \\ 0 & 0 & [T_{i3}] & 0 \\ \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & [T_{im}] \end{bmatrix}$$

Thus it follows that $[T_i - \lambda_i I] = \bigoplus_r [T_{ir}]$. If $[T_i]$ is the matrix representation of T_i with respect to the special basis for U_i, then

$$[T_i] - \lambda_i I = [T_i - \lambda_i I] = \bigoplus_r [T_{ir}]$$

Thus
$$[T_i] = \lambda_i I + \bigoplus_r [T_{ir}]$$

Substituting for $[T_{ir}]$,

$$[T_i] = \lambda_i I + \bigoplus_{r,s} \begin{bmatrix} [Q_{r1}] & 0 & 0 & & 0 \\ 0 & [Q_{r2}] & 0 & & 0 \\ 0 & 0 & [Q_{r3}] & & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & & [Q_{rs}] \end{bmatrix}$$

Let
$$[J_{rt}] = \lambda_i I + [Q_{rt}] \qquad t = 1, 2, \ldots, s$$

so that,

$$[J_{rt}] = \begin{bmatrix} \lambda_i & 1 & 0 & 0 & \cdots & 0 \\ 0 & \lambda_i & 1 & 0 & \cdots & 0 \\ 0 & 0 & \lambda_i & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & & \vdots \\ \vdots & \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & 0 & \cdots & 1 \\ 0 & 0 & 0 & 0 & \cdots & \lambda_i \end{bmatrix}$$

The size of each $[J_{rt}]$, $t = 1, 2, \ldots, s$, is the same for a given $[T_{ir}]$.

The $[J_{rt}]$s are called **Jordan blocks**.

Thus if $[T_i]$ is the matrix representation of T_i on U_i with respect to the specially constructed basis for U_i,

$$[T_i] = \begin{bmatrix} [J_1] & 0 & 0 & 0 \\ 0 & [J_2] & 0 & 0 \\ 0 & 0 & [J_3] & 0 \\ \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & [J_m] \end{bmatrix}$$

where each $[J_r]$ is of the form

$$[J_r] = \begin{bmatrix} [J_{r1}] & 0 & 0 & 0 \\ 0 & [J_{r2}] & 0 & 0 \\ 0 & 0 & [J_{r3}] & 0 \\ \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & [J_{rs}] \end{bmatrix}$$

for $r = 1, 2, \ldots, m$.

Observe that **m** will indicate the **number of Jordan blocks** of size **s** corresponding to the eigenvalue λ_i.

We draw the following conclusions to arrive at the Jordan form of the linear operator T on the vector space V :

i) Since $V = U_1 \oplus U_2 \oplus \ldots \oplus U_k$, the union of the special bases for U_i, $i = 1, 2, \ldots, k$, is a basis for V. (It is **the** special basis for V).

ii) Since $\mathcal{T}_i$ is the restriction of $\mathcal{T}$ to U_i, $i = 1, 2, \ldots, k$, we get

$$\mathcal{T} = \mathcal{T}_1 \oplus \mathcal{T}_2 \oplus \ldots \oplus \mathcal{T}_k$$

by the definition (p.13) of the reduction of $\mathcal{T}$ by $(U_1, U_2, \ldots, U_k)$.

Thus the matrix representation of $\mathcal{T}$ with respect to the special basis for V, will be given by [T] where

$$[T] = \begin{bmatrix} [T_1] & 0 & 0 & 0 \\ 0 & [T_2] & 0 & 0 \\ 0 & 0 & [T_3] & 0 \\ \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & [T_k] \end{bmatrix}$$

This is a Jordan Canonical Form of the matrix representation of a linear operator $\mathcal{T}$ on a finite-dimensional vector space V, with respect to the special basis constructed for V □

We emphasize that the matrix [T] above is not the only Jordan form, but depends on the order in which the eigenspaces of V are chosen. However, it is possible to determine the number of Jordan blocks of a given size in each eigenspace of V [1]. These Jordan blocks are unique to $\mathcal{T}$. Different combinations of these blocks will produce different, but equally valid Jordan representations of $\mathcal{T}$.

We write the major result of this chapter as the following theorem:

Theorem 10

Let $\mathcal{T}$ be a linear operator on the finite-dimensional vector space V whose characteristic polynomial has dim(V) roots (not necessarily distinct) in the field $\mathbb{F}$ over which V is defined. Then

i) There exists a special basis for V such that the matrix representation of $\mathcal{T}$ has a direct-sum decomposition into Jordan block matrices.

ii) In this decomposition the number of Jordan blocks and their sizes are uniquely determined by $\mathcal{T}$ ■

Before concluding, we make the following observations concerning the minimum and characteristic polynomials of $\mathcal{T}$. First, the multiplicity of a root of the minimum polynomial (r_i in Theorem 9) is equal to the size of the largest "Jordan block" corresponding to that root. (Thus a Jordan canonical form contains only one Jordan block for each eigenvalue λ_i if and only if the minimum polynomial of $\mathcal{T}$ is equal to the characteristic polynomial of $\mathcal{T}$). Second, the characteristic and minimum polynomials do not necessarily determine a Jordan form of $\mathcal{T}$.

We conclude our discussion of the Jordan Canonical Form with some examples which illustrate the idea of the "Jordan blocks" and their relation to the characteristic and minimum polynomials.

Example 1

Let $c(\lambda) = (\lambda - 2)^3 (\lambda - 3)^2$ and $m(T) = (T - 2)^2 (T - 3)$ be the characteristic and minimum polynomials respectively of a linear operator T on the vector space V. $c(\lambda)$ indicates that the diagonal entries (eigenvalues) are 2, 2, 2, 3, and 3. $m(T)$ indicates that the size of the largest Jordan block with eigenvalue 2 is 2, and the size of the largest Jordan block with eigenvalue 3 is 1. Thus a Jordan form of T is

$$[T] = \begin{bmatrix} 2 & 1 & & & \\ 0 & 2 & & & \\ & & 2 & & \\ & & & 3 & \\ & & & & 3 \end{bmatrix}$$

Example 2

Let $c(\lambda) = (\lambda - 2)^4 (\lambda - i)^3 (\lambda + i)$ and $m(T) = (T - 2)^2 (T - i)^2 (T + i)$. Then there are two possible sets of Jordan blocks in a Jordan form of T.

$$[T] = \begin{bmatrix} 2 & 1 & & & & & & \\ 0 & 2 & & & & & & \\ & & 2 & 1 & & & & \\ & & 0 & 2 & & & & \\ & & & & i & 1 & & \\ & & & & 0 & i & & \\ & & & & & & i & \\ & & & & & & & -i \end{bmatrix} \text{ or } [T] = \begin{bmatrix} 2 & 1 & & & & & & \\ 0 & 2 & & & & & & \\ & & 2 & & & & & \\ & & & 2 & & & & \\ & & & & i & 1 & & \\ & & & & 0 & i & & \\ & & & & & & i & \\ & & & & & & & -i \end{bmatrix}$$

Example 3

Let $T: \mathbb{C}^9 \dashrightarrow \mathbb{C}^9$ be the linear operator represented by the Jordan form matrix

$$[T] = \begin{bmatrix} 1-i & 1 & & & & & & & \\ 0 & 1-i & & & & & & & \\ & & 1-i & 1 & & & & & \\ & & 0 & 1-i & & & & & \\ & & & & 3 & 1 & 0 & & \\ & & & & 0 & 3 & 1 & & \\ & & & & 0 & 0 & 3 & & \\ & & & & & & & 3 & \\ & & & & & & & & 3 \end{bmatrix}$$

Then $c(\lambda) = (\lambda + i - 1)^4 (\lambda - 3)^5$ and $m(T) = (T + i - 1)^2 (T - 3)^3$.

The Spectral Theorem

for Normal Operators

Introduction

The discussion of the Jordan form of a linear operator indicated that on a finite-dimensional space, the spectrum of a linear operator consists of its eigenvalues. However, the Jordan form is a spectral theorem limited to a finite-dimensional space and does not generalize to infinite-dimensional space.

In this chapter we shall prove a spectral theorem for normal operators on a finite-dimensional space. The proof will be related to our discussion of the Jordan form. However, unlike the Jordan form, the spectral theorem for normal operators does generalize to infinite-dimensional space.

THE BASIC DEFINITIONS AND THEOREMS

The following concepts and related theorems are necessary for our discussion. We do not discuss or prove them in detail as our aim is to relate the spectral theorem for normal operators to our discussion of the Jordan form.

An **inner product** in a complex (real) vector space H is a complex (real) valued function, denoted **(,)** whose domain is H x H such that

i) $(x, y) = \overline{(y, x)}$ where $\overline{(y, x)}$ denotes the complex conjugate of (y, x)

ii) $(\alpha_1 x_1 + \alpha_2 x_2, y) = \alpha_1(x_1, y) + \alpha_2(x_2, y)$

and for any complex (real) α_1, α_2,

iii) $(x, x) \geq 0$ and $(x, x) = 0$ iff $x = 0$

An **inner-product space** is a vector space equipped with an inner product. A real inner-product space is a **Euclidean space**. If the scalar field is complex, then an inner-product space is called a **unitary space**.

T^* is an **adjoint** of a linear operator T on an inner-product space H, if T^* is an operator on H such that $(T(x), y) = (x, T^*(y))$, for all $x, y \in$ H.

T is said to be a **self-adjoint operator** (or Hermitian) if $T = T^*$.

An operator U on a unitary space is a **unitary operator** iff $U^* = U^{-1}$.

If H is an inner-product space and U is a subspace of H, then

$$U^{\perp} = \{ y : y \in H,\ (x, y) = 0 \text{ for any } x \in U \}$$

is a subspace of H and $U^{\perp}$ is called the **orthogonal complement** of U.

A projection E on U along $U^{\perp}$ is called the **orthogonal projection** on U.

$\| E \|$ is the number by which E multiplies the length of a vector.

Theorem 1

An operator E on the inner-product space H is an orthogonal projection iff $E = E^2 = E^*$. If E is a nonzero orthogonal projection, then $\| E \| = 1$ ■

An operator T on the inner-product space H is said to be a **normal operator** if and only if $TT^* = T^*T$. (Both self-adjoint and unitary operators are normal operators since they both commute with their adjoint).

Theorem 2

If T is a normal operator on an inner-product space H, then x is an eigenvector of T corresponding to the eigenvalue λ if and only if x is an eigenvector of T^* corresponding to the eigenvalue $\bar{\lambda}$ ■

Theorem 3

It follows from the above theorem that if x is in the null space of T, then x is in the null space of T^* ■

The spectral theorem for normal operators

There are several possible proofs of the spectral theorem for normal operators using the properties of orthogonal projections. However, as mentioned at the outset, the proof we present is related to the proof of the existence of a Jordan form for linear operators on a finite-dimensional space. The generalization to infinite-dimensions requires different techniques.

Theorem 4

Let T be a normal operator on the finite-dimensional inner-product space H. If $\lambda_1, \lambda_2, \ldots, \lambda_k$, are the distinct eigenvalues of T, then there exist orthogonal projections $E_1, E_2, \ldots, E_k$, such that E_i is a projection on the eigenspace corresponding to λ_i, $i = 1, 2, \ldots, k$.

Furthermore,

$$\text{i)} \quad T = \sum_{i=1}^{k} \lambda_i E_i$$

$$\text{ii)} \quad \sum_{i=1}^{k} E_i = I$$

$$\text{iii)} \quad E_i E_j = 0, \qquad i \neq j$$

Proof

Let $\varpi(t) = (t - \lambda_1)(t - \lambda_2) \ldots (t - \lambda_k)$, correspond to $\varpi(T)$, the minimum polynomial of T.

If $k < 2$, then either $\varpi(T)$ is of the form $T + \alpha I = 0$, or $T = 0$. In either case the proof of the theorem is trivial. We prove that the theorem holds true for $k \geq 2$.

Define $p_i(t)$ by

$$p_i(t) = \prod_{\substack{j=i \\ j \neq i}}^{k} \frac{(t - \lambda_j)}{(\lambda_i - \lambda_j)}$$

Then

$$p_1(t) = \frac{(t - \lambda_2)\ (t - \lambda_3)\ \ldots\ (t - \lambda_k)}{(\lambda_1 - \lambda_2)\,(\lambda_1 - \lambda_3)\ \ldots\ (\lambda_1 - \lambda_k)}$$

$$p_2(t) = \frac{(t - \lambda_1)\ (t - \lambda_3)\ \ldots\ (t - \lambda_k)}{(\lambda_2 - \lambda_1)\,(\lambda_2 - \lambda_3)\ \ldots\ (\lambda_2 - \lambda_k)}$$

$$\vdots$$

Thus $p_i(t)$ has degree $k - 1$ for all $i = 1, 2, \ldots, k$.

Let $$p(t) = \sum_{i=1}^{k} p_i(t)$$

Then $p(t)$ has degree $k - 1$.

Notice that $p_1(\lambda_1) = 1$, but $p_2(\lambda_1) = p_3(\lambda_1) = \ldots = p_k(\lambda_1) = 0$.

Also $p_2(\lambda_2) = 1$, but $p_1(\lambda_2) = p_3(\lambda_2) = \ldots = p_k(\lambda_2) = 0$.

and in general, $p_i(\lambda_j) = \delta_{ij}$.

Suppose $s(t)$ is a polynomial of degree $k - 1$ or less. Then we claim that $s(t)$ can be expressed as

$$s(t) = \sum_{i=1}^{k} s(\lambda_i)\, p_i(t)$$

This is because

$$s(\lambda_j) - \sum_{i=1}^{k} s(\lambda_i)\, p_i(\lambda_j) = s(\lambda_j) - s(\lambda_j)\, p_j(\lambda_j) = s(\lambda_j) - s(\lambda_j) = 0$$

since $p_i(\lambda_j) = \delta_{ij}$ for $j = 1, 2, \ldots, k$.

Then $s(t) - \sum_{i=1}^{k} s(\lambda_i)\, p_i(t)$ is a polynomial of degree $\leq k - 1$ with k distinct zeros ($t = \lambda_1, \lambda_2, \ldots, \lambda_k$). Thus the polynomial must be zero. We consider two possibilities for $s(t)$:

<u>Case 1</u> : Let $s(t) = 1$. Then $s(\lambda_i) = 1$. Thus $\sum_{i=1}^{k} p_i(t) = 1$

<u>Case 2</u> : Let $s(t) = t$. Then $s(\lambda_i) = \lambda_i$. Thus $\sum_{i=1}^{k} \lambda_i\, p_i(t) = t$

Also observe that for $i \neq j$,

$$p_i(t)\, p_j(t) = \prod_{\substack{j=1 \\ j \neq i}}^{k} \frac{(t - \lambda_j)}{(\lambda_i - \lambda_j)} \prod_{\substack{i=1 \\ i \neq j}}^{k} \frac{(t - \lambda_i)}{(\lambda_j - \lambda_i)}$$

Therefore $p_i(t)\, p_j(t) = \varpi(t)\, q(t)$, for $i \neq j$, where $q(t)$ is a polynomial in t .

Then since $\varpi(t) = 0$,

$$p_i(t)\, p_j(t) = 0 \qquad \text{for } i \neq j$$

Thus we have

i) $$\sum_{i=1}^{k} p_i(t) = 1$$

ii) $$\sum_{i=1}^{k} \lambda_i\, p_i(t) = t$$

iii) $$p_i(t)\, p_j(t) = 0 \qquad \text{for } i \neq j$$

Now the set of all polynomials in t over the scalar field is a ring, say $F(t)$. Similarly, the set of all polynomials in T over the scalar field is a ring, say $F(T)$. There is thus an isomorphic mapping from $F(t)$ to $F(T)$ which preserves polynomial identities.

We can therefore rewrite the results obtained above as,

i) $$\sum_{i=1}^{k} p_i(T) = I$$

ii) $$\sum_{i=1}^{k} \lambda_i\, p_i(T) = T$$

iii) $$p_i(T)\, p_j(T) = 0 \qquad \text{for } i \neq j$$

We let $E_j = p_j(T)$ $j = 1, 2, \ldots, k$. Then it follows that the three conditions

$$\text{i)} \quad T = \sum_{i=1}^{k} \lambda_i E_i$$

$$\text{ii)} \quad \sum_{i=1}^{k} E_i = I$$

$$\text{iii)} \quad E_i E_j = 0, \qquad \text{for } i \neq j$$

are satisfied. To complete the proof, we must show that $E_j = p_j(T)$ is an orthogonal projection on the eigenspace of λ_j, for $j = 1, 2, \ldots, k$.

Now $$I = \sum_{i=1}^{k} E_i$$

Then $$E_j = \sum_{i=1}^{k} E_i E_j = E_j^2 \qquad (\text{since } E_i E_j = 0,\ i \neq j)$$

Therefore E_j is a projection, $j = 1, 2, \ldots, k$.

Let $x \in H$. Then

$$E_j(x - E_j(x)) = E_j(x) - E_j^2(x) = E_j(x) - E_j(x) = 0$$

Thus $x \in N_{E_j}$.

Then from Theorem 3, $x \in N_{E_j^*}$.

Therefore $$E_j^*(x - E_j(x)) = 0$$

or $$E_j^*(x) - E_j^* E_j(x) = 0 \qquad \text{for all } x \in H.$$

Then $E_j^* = E_j^* E_j$. Thus

$$E_j = E_j^{**} = (E_j^* E_j)^* = E_j^* E_j = E_j^*$$

Therefore E_j is an orthogonal projection for each $j = 1, 2, \ldots, k$.

Finally, we need to prove that E_j projects onto the eigenspace of λ_j.

Suppose x is an eigenvector of T corresponding to the eigenvalue λ_j. Then

since $\sum_{i=1}^{k} \lambda_i E_i = T$,

$$\sum_{i=1}^{k} \lambda_i E_i(x) = T(x) = \lambda_j x$$

Also since $I = \sum_{i=1}^{k} E_i$,

$$x = \sum_{i=1}^{k} E_i(x)$$

Thus
$$\sum_{i=1}^{k} \lambda_i E_i(x) = \lambda_j \sum_{i=1}^{k} E_i(x)$$

Therefore
$$\sum_{i=1}^{k} (\lambda_i - \lambda_j) E_i(x) = 0$$

Then
$$\| \sum_{i=1}^{k} (\lambda_i - \lambda_j) E_i(x) \|^2 = 0$$

$$\therefore \|(\lambda_1 - \lambda_j)E_1(x) + (\lambda_2 - \lambda_j)E_2(x) + \ldots + (\lambda_j - \lambda_j)E_j(x) + \ldots + (\lambda_k - \lambda_j)E_k(x)\|^2 = 0$$

Then since $\mathcal{E}_i\mathcal{E}_j = 0$, for $i \neq j$,

$$\sum_{i=1}^{k} (\lambda_i - \lambda_j)^2 \| \mathcal{E}_i(x) \|^2 = 0$$

But for the sum of non-negative quantities to be zero, each summand must be zero.

Thus for $i \neq j$, $\qquad \mathcal{E}_i(x) = 0$

Then since $x = \sum_{i=1}^{k} \mathcal{E}_i(x)$,

$$\mathcal{E}_j(x) = x$$

On the other hand, if $\mathcal{E}_j(x) = x$, then $\mathcal{T}(x) = \mathcal{T}\mathcal{E}_j(x)$. Also since $\mathcal{E}_j = p_j(\mathcal{T})$ is a polynomial in $\mathcal{T}$, it commutes with $\mathcal{T}$.

Thus $\qquad \mathcal{T}(x) = \mathcal{T}\mathcal{E}_j(x) = \mathcal{E}_j\mathcal{T}(x) = \mathcal{E}_j\left(\sum_{i=1}^{k} \lambda_i \mathcal{E}_i(x)\right)$

$$= \lambda_j \mathcal{E}_j\mathcal{E}_j(x) \qquad (\mathcal{E}_j\mathcal{E}_i = 0,\ i \neq j)$$

$$= \lambda_j \mathcal{E}_j(x) \qquad (\mathcal{E}_j(x) = x)$$

$$\therefore\ \mathcal{T}(x) = \lambda_j x$$

Therefore x is an eigenvector corresponding to the eigenvalue λ_j, and the range of $\mathcal{E}_j$ is the set of eigenvectors of $\mathcal{T}$ corresponding to λ_j ■

Conclusion

Quantum mechanics is one area in modern physics where the mathematical theory of linear algebra and inner product spaces is used. Damiano and Little [2] write that

> In the nineteenth century it was discovered that when atoms of a substance (e.g., gases such as hydrogen) are excited, they give off light of certain precisely defined wavelengths (colors). This is what happens, for example, in "neon" signs. An electric current is passed through a glass tube containing a gas. The excited atoms of the gas give off light of the desired color. In fact, the wavelengths (or "spectral lines") of the light emitted by a glowing object are so characteristic of the elements present that they can be used to determine the composition of the object. There was no systematic explanation for this behavior available at the time it was first noticed, however.
>
> Toward the end of the nineteenth century, mathematicians studying analogs of our symmetric linear mappings on function spaces noted a similarity between the discrete spectral lines of atoms and the pattern distribution of eigenvalues of these linear mappings. Perhaps motivated by this analogy, the German mathematician Hilbert, who proved the first general version of the theorem in question (but in the function space setup), called his result the "spectral theorem."
>
> In the 1920s and 1930s the new physical theory of quantum mechanics was developed to explain phenomena such as the spectral emission lines mentioned before, and in a rather amazing vindication of the earlier mathematicians' intuition, it was found that the wavelengths of the emission spectrum of an atom were indeed related to the eigenvalues of a certain linear operator. This was the Schrödinger operator that appears in the basic differential equation of quantum mechanics. pp 206 - 207

The concept of a spectral theorem for a given operator is thus applicable to the study of external phenomena.

NOTES

[1] In his Linear Algebra (New York : Harcourt Brace, 1976), Michael O'Nan gives the following formula for determining the exact numbers and sizes of Jordan blocks for a given linear operator.

> Let T be an operator on the finite-dimensional vector space V and λ be an eigenvalue of T. Then in the Jordan normal form of T, the number of Jordan blocks of size k and diagonal entry λ is
>
> $$r([T-\lambda I]^{k-1}) - 2r([T-\lambda I]^{k}) + r([T-\lambda I]^{k+1})$$
>
> p. 315

N.B. $r([T])$ is the rank of $[T]$.

[2] A Course in Linear Algebra, David B. Damiano and John B. Little, (New York : Harcourt Brace, 1988).

BIBLIOGRAPHY

Damiano, David, and John B. Little. A Course in Linear Algebra. New York : Harcourt Brace, 1988.

O'Nan, Michael. Linear Algebra. New York : Harcourt Brace, 1976.

Scroggs, James E. Linear Algebra. Belmont : Brooks, 1970.

DATE DUE

GAYLORD

PRINTED IN U.S.A.

Archives
512.8-D457t
De Silva, Rapti Manohara
AUTHOR
Two spectral theorems
TITLE

Archives
512.8-D457t
De Silva, Rapti Manohara

Sweet Briar College Library
Sweet Briar, VA 24595